DOCUMENT RESUME

ED 136 780 IR 004 567

AUTHOR Levin, James A.; Archbold, Armar A.

TITLE Working Papers in Dialogue Modeling, Volume 1.
INSTITUTION University of Southern California, Marina del Rey. Information Sciences Inst.
SPONS AGENCY Advanced Research Projects Agency (DOD), Washington, D.C.; Office of Naval Research, Arlington, Va. Personnel and Training Research Programs Office.
REPORT NO ISI/RR-77-55
PUB DATE Jan 77
CONTRACT N00014-75-C-0710; NR-154-374
NOTE 82p.; For Volume 2, see IR 004 568

EDRS PRICE MF-$0.83 HC-$4.67 Plus Postage.
DESCRIPTORS *Artificial Intelligence; Communication (Thought Transfer); *Computational Linguistics; Computer Oriented Programs; Computers; Dialogue; *Language Research; Man Machine Systems; *Models; Programing Languages
IDENTIFIERS Dialogue Modeling

ABSTRACT

The five technical working papers that comprise the two volumes of this document are related to the problem of creating a valid process model of human communication in dialogue. In Volume 1 both papers consider reference as a phenomenon in text. The first surveys reference identification and resolution methods in various existing natural language processors. The second paper explores the broader problem of reference, focusing on text reference and propositional reference. It develops problems and proposals for defining these categories of reference phenomena, and for detecting their presence. (Author/WBC)

****
* Documents acquired by ERIC include many informal unpublished *
* materials not available from other sources. ERIC makes every effort *
* to obtain the best copy available. Nevertheless, items of marginal *
* reproducibility are often encountered and this affects the quality *
* of the microfiche and hardcopy reproductions ERIC makes available *
* via the ERIC Document Reproduction Service (EDRS). EDRS is not *
* responsible for the quality of the original document. Reproductions *
* supplied by EDRS are the best that can be made from the original. *
****

ED136780

ARPA ORDER NO. 2930
NR 154-374

U.S. DEPARTMENT OF HEALTH
EDUCATION & WELFARE
NATIONAL INSTITUTE OF
EDUCATION

THIS DOCUMENT HAS BEEN REPRODUCED EXACTLY AS RECEIVED FROM THE PERSON OR ORGANIZATION ORIGINATING IT. POINTS OF VIEW OR OPINIONS STATED DO NOT NECESSARILY REPRESENT OFFICIAL NATIONAL INSTITUTE OF EDUCATION POSITION OR POLICY

ISI/RR-77-55
January 1977

James A. Levin

Armar A. Archbold

Working Papers in Dialogue Modeling, Volume 1

IR004567

2

INFORMATION SCIENCES INSTITUTE

UNIVERSITY OF SOUTHERN CALIFORNIA

4676 Admiralty Way/Marina del Rey/California 90291
(213) 822-1511

Preparation of this paper was supported by the Office of Naval Research, Personnel and Training Research Programs, Code 458, under Contract N00014-75-C-0710, NR 154-374, under terms of ARPA Order Number 2930.
The views and conclusions contained in this document are those of the author(s) and should not be interpreted as necessarily representing the official policies, either expressed or implied of the Office of Naval Research, the Defense Advanced Research Projects Agency, or the U.S. Government.
This document is approved for public release and sale; distribution is unlimited.

UNCLASSIFIED

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

REPORT DOCUMENTATION PAGE		READ INSTRUCTIONS BEFORE COMPLETING FORM
1. REPORT NUMBER ISI/RR-77-55	2. GOVT ACCESSION NO.	3. RECIPIENT'S CATALOG NUMBER
4. TITLE (and Subtitle) WORKING PAPERS IN DIALOGUE MODELING - Volume 1		5. TYPE OF REPORT & PERIOD COVERED Research 6. PERFORMING ORG. REPORT NUMBER
7. AUTHOR(s) James A. Levin, Armar A. Archbold		8. CONTRACT OR GRANT NUMBER(s) N00014-75-C-0710
9. PERFORMING ORGANIZATION NAME AND ADDRESS USC/Information Sciences Institute 4676 Admiralty Way Marina del Rey, California 90291		10. PROGRAM ELEMENT, PROJECT, TASK AREA & WORK UNIT NUMBERS 61153N RR042-06-01 RR042-06 NR154-374
11. CONTROLLING OFFICE NAME AND ADDRESS Cybernetics Technology Office Advanced Research Projects Agency 1400 Wilson Blvd., Arlington, VA 22209		12. REPORT DATE January 1977 13. NUMBER OF PAGES 80
14. MONITORING AGENCY NAME & ADDRESS(if different from Controlling Office) Personnel and Training Research Programs Office of Naval Research - Code 458 800 No. Quincy St. Arlington, VA 22217		15. SECURITY CLASS. (of this report) UNCLASSIFIED 15a. DECLASSIFICATION/DOWNGRADING SCHEDULE
16. DISTRIBUTION STATEMENT (of this Report) Approved for public release; distribution unlimited		
17. DISTRIBUTION STATEMENT (of the abstract entered in Block 20, if different from Report)		
18. SUPPLEMENTARY NOTES		
19. KEY WORDS (Continue on reverse side if necessary and identify by block number) artificial intelligence, cognitive psychology, computer dialogue, linguistic comprehension, process model, natural language processing, reference.		
20. ABSTRACT (Continue on reverse side if necessary and identify by block number) (OVER)		

DD FORM 1 JAN 73 1473 EDITION OF 1 NOV 65 IS OBSOLETE
S/N 0102-014-6601

UNCLASSIFIED

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

20. ABSTRACT

The five technical working papers that compose this document (which appears in two volumes) were prepared as part of the Dialogue Modeling Project at ISI. Though diverse in scope, all are related to the problem of creating a valid process model of human communication in dialogue. All are unpublished and all but one are in a form intended for internal use by the project team; however, they are of interest beyond the boundaries of the project and have implications for related work in modeling human communication.

In Volume 1 both papers are on reference as a phenomenon in text. The first surveys reference identification and resolution methods in various existing natural language processors. The other paper explores the broader problem of reference, focusing on text reference and propositional reference. It develops problems and proposals for defining these categories of reference phenomena and for detecting instances of them.

In Volume 2 the first paper concerns study methodology. It raises some of the following issues: how to choose between system-building and process-building, why studying cases is preferable to implementing general language-use functions as programs, how to control ad-hocness of results, why it is important to orient toward communication phenomena (in contrast to form phenomena) when studying natural language. The second is a design paper on the Match process of the Dialogue Model System, exploring methods for making it efficient and selective in its actions. The third concerns the structure of persuasion dialogues, in particular how justification of actions appears in argumentation.

Each working paper appears with its original abstract or introduction.

ARPA ORDER NO. 2930
NR 154-374

U S DEPARTMENT OF HEALTH,
EDUCATION & WELFARE
NATIONAL INSTITUTE OF
EDUCATION

THIS DOCUMENT HAS BEEN REPRODUCED EXACTLY AS RECEIVED FROM THE PERSON OR ORGANIZATION ORIGINATING IT. POINTS OF VIEW OR OPINIONS STATED DO NOT NECESSARILY REPRESENT OFFICIAL NATIONAL INSTITUTE OF EDUCATION POSITION OR POLICY

ISI/RR-77-55
January 1977

James A. Levin
Armar A. Archbold

Working Papers in Dialogue Modeling, Volume 1

INFORMATION SCIENCES INSTITUTE
UNIVERSITY OF SOUTHERN CALIFORNIA
4676 Admiralty Way/Marina del Rey/California 90291
(213) 822-1511

Preparation of this paper was supported by the Office of Naval Research, Personnel and Training Research Programs, Code 458, under Contract N00014-75-C-0710, NR 154-374, under terms of ARPA Order Number 2930.
The views and conclusions contained in this document are those of the author(s) and should not be interpreted as necessarily representing the official policies, either expressed or implied, of the Office of Naval Research, the Defense Advanced Research Projects Agency, or the U.S. Government.
This document is approved for public release and sale; distribution is unlimited.

TABLE OF CONTENTS

	Page
Volume 1	
Section 1 - Process Models of Reference by James A. Levin	1.1
Section 2 - Text Reference and Repeated Prepositional Reference: Concepts and Detection Procedures by Armar A. Archbold	1.2
Volume 2	
Section 1 - Improving Methodology in Natural Language Processing by William C. Mann	2.1
Section 2 - A Selector Routine to Generate Candidates for Match by Greg Scragg	2.2
Section 3 - A Study of Some Argument-forms in Persuasion-Dialogue by Armar A. Archbold	2.3

ABSTRACT

The five technical working papers that compose this document (which appears in two volumes) were prepared as part of the Dialogue Modeling Project at ISI. Though diverse in scope, all are related to the problem of creating a valid process model of human communication in dialogue. All are unpublished and all but one are in a form intended for internal use by the project team; however, they are of interest beyond the boundaries of the project and have implications for related work in modeling human communication.

In Volume 1 both papers are on reference as a phenomenon in text. The first surveys reference identification and resolution methods in various existing natural language processors. The other paper explores the broader problem of reference, focusing on text reference and propositional reference. It develops problems and proposals for defining these categories of reference phenomena and for detecting instances of them.

In Volume 2 the first paper concerns study methodology. It raises some of the following issues: how to choose between system-building and process-building, why studying cases is preferable to implementing general language-use functions as programs, how to control ad-hocness of results, why it is important to orient toward communication phenomena (in contrast to form phenomena) when studying natural language. The second is a design paper on the Match process of the Dialogue Model System, exploring methods for making it efficient and selective in its actions. The third concerns the structure of persuasion dialogues, in particular how justification of actions appears in argumentation.

Each working paper appears with its original abstract or introduction.

Working Papers in Dialogue Modeling

Volume 1 - Section 1

Process Models of Reference

James A. Levin

Original Date: November 20, 1975

Table of Contents

1. Repeated Reference	2
2. Repeated Reference in Existing Systems	3
2.1 Repeated Reference in Verbworld	6
2.2 Repeated Reference in SHRDLU	8
2.3 Repeated Reference in Semantically Oriented Systems	9
3. Classification of the Bag of Repeated Reference Tricks	10
3.1 The Possible Co-referential Expressions	14
3.2 Aspects of the Current Expression	14
4. An Activation Model of Repeated Reference	14
4.1 Rules for Modifying the Activation Rates of Possible referents	15
4.2 Rules for Specifying the Referent	16
5. Complex Repeated Reference	17
5.1 Complex Repeated Reference and the Activation Model	18
5.2 Quotes	18
6. Non-Repeated Reference	19
6.1 Foreground Reference and the Activation Model	19
6.2 Failure to Find a Referent	20
7. Referents and Reference	20
8. Summary	21

One of the challenges in modeling human language capability is the fact that words are used to convey different meanings in different contexts. For example, the meaning of a pronoun is heavily dependent on the context in which it occurs, since it is used to refer to some entity mentioned elsewhere and has the same referent as this other expression. The meaning of instances of "it" may change radically from occurence to occurence, even within the same utterance. In the following portion of dialogue * :

"O: Did Runoff produce any output?
L: I dont know but I thought *it* would ask me what name *it* would be."

In the second utterance, the two "it"s refer to completely different things expressed in the first utterance, the first to "Runoff" and the second to "output".

Many of the process models of language understanding have dealt with this issue of repeated reference, in which one expression (often containing a pronoun or a determiner) derives meaning from some other expression in the discourse. Generally, these models contain a set of heuristic rules for ordering possible co-referential expressions and for testing them for appropriateness. For these models, the rules form a "bag of tricks" approach to repeated reference, rather than a systematic theory.

In this paper, we will examine in detail several models of language understanding that deal with repeated reference. After examining the operation of these models on a simple dialogue, we will classify the assembled "bag of repeated reference tricks" into two categories, and from this classification, propose a new approach for the process modeling of repeated reference.

1. Repeated Reference

Repeated Reference occurs when two sets of words in a discourse refer to the same concept. Let us call a set of words in an utterance defined as a unit on syntactic grounds an Expression. A Co-referential Expression is an Expression that has the same referent as some other Expression. Not all Expressions are Co-referential Expressions. Sometimes definite noun phrases are used to refer to generic concepts ("Do you know how *the mail system* works?"). Other times, an Expression can introduce a unique concept ("...*the ISI line printer*..."), or a concept that is unique in the given situation ("I just used a system called XOFF and it didn't give me *the normal output*"). Other Expressions in fact don't refer to anything ("*It's* 5 oclock.").

It is important to distinguish between Expressions and Concepts. Expressions are sets of words, while Concepts are the abstract entities which are the referents of Expressions (as well as referents of other non-verbal stimuli).

---

* Unless otherwise specified, all examples in this paper are taken from naturally occurring dialogues collected by the Dialogue Modeling Project at ISI. These dialogues are between a computer operator (labeled "O") and a computer user ("L"). The participants communicated remotely by typing into computer terminals using the TENEX "link" facility: whatever either person types appears simultaneously on both terminals.

We can easily detect Expressions because they are defined on a syntactic basis. In fact, we could use one of the existing parsing systems to mechanically detect Expressions. However, this isn't true for Co-referential Expressions. Some people (Baranofsky, 1970; Olney, 1969) have investigated the detection of Co-referential Expressions (sometimes called "anaphoric expressions") using certain syntactic aspects to distinguish the Co-referential Expressions from non-repeated ones. The language understanding systems described here all approach this detection problem by delaying the decision until it is trivial. They all look for preexisting referents for all Expressions, and those that have referents that are also the referents of other Expressions are then Co-referential Expressions.

2. Repeated Reference in Existing Systems

We will now concentrate on the heuristics for finding an existing referent, given an Expression. The operation of some of these will be illustrated on a simple dialogue, shown in figure 1. This is a real dialogue between a computer user and a computer operator typing over computer terminals using the "lin[illegible]" facility.

```
    LINK FROM [L], TTY 42

1.   1  2  3  4    5    6   7    8  9    10   11 12  13  14   15
L:  HOW DO I GET RUNOFF TO WORK, I KEEP XEQTN IT BUT IT JUST GRABS

1.  16  17    18  19   20   21   22  23   24   25 26   27     28
    MY INPUT FILE AND THEN SAYS DONE BUT GIVES ME NO OUTPUT?  GA


2.   1    2      3    4  5   6   7     8
O:  THE OUTPUT COMES OUT ON THE LINE PRINTER


3.    1   2   3    4   5  6  7  8  9   10   11 12 13    14
L:  THROW IT AWAY BUT CAN I GET IT TO GE\EO TO A FILE?  GA


4.     1     2     3      4   5   6    7    8    9    10    11    12
O:  CONFIRM YOUR COMMANDS WITH A COMMA AND YOU'LL BE QUERIED FOR FILES,

4.  13    14
    ETC.  GA


5.   1    2
L:  THX MUCHO


    BREAK (LINKS)
```

Figure 1: Dialogue between a TENEX Operator ("O") and a User ("L")
(The numbers and speaker labels have been added).

Our Dialogue Modeling Project at ISI has developed a procedure to gather annotations from observers of a dialogue transcript to use in evaluating dialogue models. Figure 2 shows the annotation produced by one observer of repeated reference phenomena in the dialogue in figure 1, using our repeated reference instructions (given in Mann, Moore, Levin & Carlisle, 1975). We will use these observations to evaluate the various repeated reference heuristics.

2.3 Repeated Reference in Verbworld

Verbworld (Rumelhart & Levin, 1975) is a recent language understanding system that deals with repeated reference in some detail.

There are two parts of this system where repeated reference is handled: the heuristics for handling pronoun reference (developed by Art Graesser) and the heuristics for handling definite determiners (developed by Donald Norman).

Pronouns

Third person pronouns (HE, SHE, IT, THEY, THEM, HIM, HER):
1. Look in the previous clauses for uses of this same pronoun.
If one is found, it is a Co-referential Expression.
else
2. Examine the noun phrases in the previous clause,
looking for a concept that matches the number and gender constraints.
Examine each clause in the following order: subject, object,
prepositional phrases. If a match is found, it is a Co-referential
Expression.
else
3. Start over with step 2.

Reflexive pronouns: (HIMSELF, HERSELF, ITSELF, THEMSELVES)
Examine previous noun phrases in this same clause for a concept with
the appropriate gender and number. If there is one, the phrase is a
Co-referential Expression.

Possessive pronouns: (HIS, HER, ITS, THEIR)
1. Examine noun phrases within the sentence, as in the case
of reflexive pronouns. If there is a match, it is a Co-referential
Expression.
else
2. Examine noun phrases in previous clauses, as in the case of third person pronouns.

Determiners

1. Evaluate the noun phrase. If the value is a unique concept, then
that value is the referent.
else
2. Evaluate any relative clauses in the noun phrase, and if this
results in an unique instance, this instance is the referent.
else

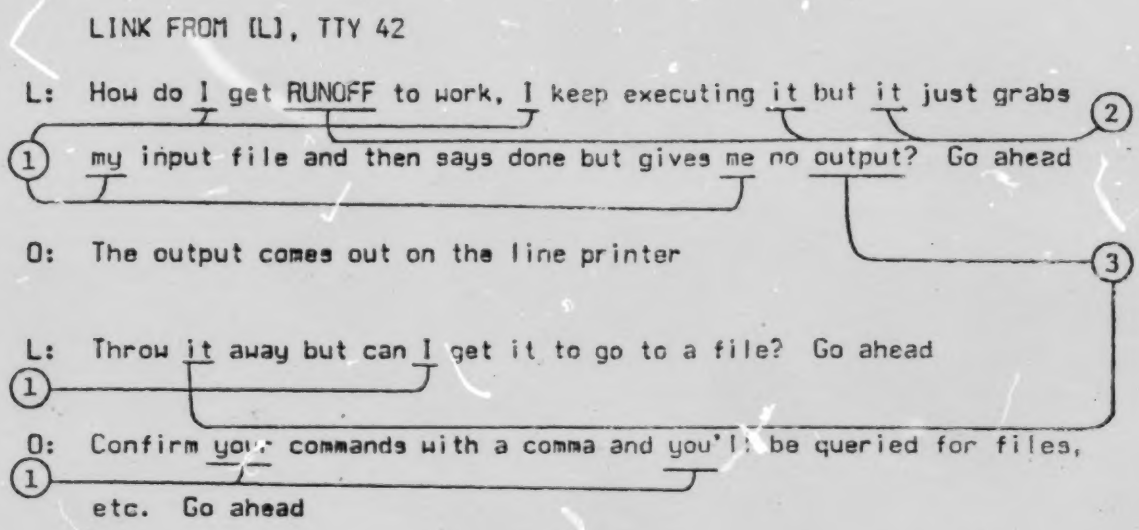
LINK FROM [L], TTY 42
L: How do I get RUNOFF to work, I keep executing it but it just grabs my input file and then says done but gives me no output? Go ahead
O: The output comes out on the line printer
L: Throw it away but can I get it to go to a file? Go ahead
O: Confirm your commands with a comma and you'll be queried for files, etc. Go ahead
L: Thanks much
BREAK (LINKS)
1
2
3

Figure 2: Dialogue annotated for Repeated Reference

3. Examine recent concepts mentioned in this and previous clauses and if there is a match with the features of the rest of the noun phrase, that concept is the referent. In particular, if an exemplar of the general concept was recently mentioned, that exemplar is the referent.

else

4. Assume that this determiner is being used in an indefinite sense, and create a new instance as the referent.

This system (like many of the others described below) was designed to be a participant in a dialogue rather than an observer of it. To enable it to operate as the observer of a dialogue, we must make one minor modification to the above heuristics: the first person pronouns ("I", "my", "me", "mine"), take as their referent the present speaker; the second person pronouns ("you", "your", "yours") take the other participant as their referent.

Let us examine a hand-simulated operation of Verbworld on the dialogue in figure 1.

It would do well on all the personal pronouns in the dialogue, using the modified rule described above, assigning as referents the participants O and L.

In handling the "it" at 1.11, Verbworld finds "Runoff" (at 1.5) as its Co-referential Expression because (1) There are no "it"s in previous clauses, (2) The subject of the previous clause "I" isn't neuter and (3) The object "Runoff" is neuter.

For the "it" at 1.13, Verbworld finds the previous "it" at 1.11 and takes it as co-referential (and therefore , RUNOFF as its referent).

Both of these assignments correspond to the annotations by the observer (figure 2).

However, in turn 2, Verbworld runs into trouble with the definite noun phrase ("the output") at 2.1-2.2.. Depending on what referent the Expression "no output" 1.26-1.27 has, the heuristics will either find no repeated reference, or, worse, will find that expression ("no output") as co-referential because (1)"output" can't be evaluated to a unique concept, (2)there are no relative clauses, and (3)if "no output" is stored as a kind of output, it will fit the constraints of the Expression.

In turn three, the pronoun heuristics also run into trouble. The "it" at 3.2 is assigned "the output" (2.1-2.2) as co-referential. However, the annotations show "the output" as the generic class, and "it" (3.2) as an instance of that class.

Furthermore, the second "it" (3.8) is assigned the first "it" (3.2) as co-referential. The observer distinguishes these as different, as we can easily see from the semantics (You can't throw something away and then get that same thing to go to a file).

Finally, the observer annotated the generic "files", at 4.12 as a superset of the indefinite "a file" at 3.13. The Verbworld system would also make this assignment, because of the way the indefinite article "a" is defined.

2.2 Repeated Reference in SHRDLU

One of the most impressive langu ge understanding systems to date is SHRDLU, developed by Winograd (1972). This system keeps track of "overall discourse", and looks in this overall discourse context for referent concepts. Winograd implemented a set of heuristics for repeated reference, most of which are summarized below:

General Tricks for Determiners

1. Find the set of all known objects that match the rest of the noun phrase.

then

2. If the right number are known, the object or the set of objects is the referent.

else

3. If there are too few objects, try to reparse the sentence (if this fails, print out a stored phrase asking for what the person meant).

or

4. If there are too many, try to find which were mentioned most recently. (if the right set of objects cannot be found, try reparsing as above, but with different error message about which were meant.)

General Tricks for Pronouns ("IT")

1. If there is another "IT" previously in the same sentence, it is a Co-referential Expression.

else

2. If there is another "IT" in the previous sentence, it is a Co-referential Expression.

else

3. For complex embedded NP's, check whether the "IT" is a reference to the NP it is in.

else

4. Look through previous clauses, looking for possible co-referential expressions, assigning plausibilities on the basis of the following:

   a. Subject > Object > Prep Phrases
   b. Main clause > Subordinate clause
   c. "Focused" objects > non-focused

Choose the most plausible matching concept as the referent.

Special Case Tricks

IT: if used as propositional reference, the most recent action mentioned by the other participant is the referent.
THAT: if used as propositional reference, the most recent action mentioned by either participant is the referent.

I: the referent is :FRIEND
YOU: the referent is :SHRDLU

Now, despite the differences in SHRDLU's heuristics from those of Verbworld, it would produce the same referent resolutions on the figure 1 dialogue, incorrectly assigning the same referent to the "it"s in turn 3 and running into problems assigning "the output" in turn 2.

Why do these systems do so poorly on this simple dialogue? The Verbworld heuristics, for example, correctly found co-referential expressions for pronouns in 90% of randomly chosen text from an encyclopedia (Rumelhart & Levin, 1975). This disparity in performance gives us a hint of what the problem is. The encyclopedia contains only well-formed grammatical sentences, while the dialogue in figure 1 contains many ill-formed utterances, as is typical of real dialogue. Since most of the heuristics discussed so far are based mostly on surface syntactic features, it's not surprising that they do poorly on the syntactically ill-formed utterances in real dialogue.

Now, this shouldn't be taken as a claim that syntactically based heuristics are useless - only that there must be additional heuristics using semantic and pragmatic features that contribute to determining that Expressions are Co-referential.

2.3 Repeated Reference in Semantically Oriented Systems

Some systems find the referents for Expressions solely on the basis of semantic features, completely ignoring syntactic aspects. Quillian (1969) described such a system, and Rieger (1974) used a similar approach to do the repeated referencing for definite noun phrases within the MARGIE System (Schank, Goldman, Rieger & Riesbeck, 1973)

In this approach, a set of specifications for the referent are collected and a match process is conducted over the set of concepts in the current context (concepts that have been recently mentioned or used recently in some inference). A concept in context that matches the specifications is then a prime candidate for being the referent of the Expression.

Let us look at how a semantically based repeated reference system would deal with the dialogue in figure 1. As with the syntactic system, it handles the personal pronouns, given the modification to make the present speaker a specification of the referent of the pronoun "I", and the other participant a specification of the referent of "you".

For the "it" at 1.11, the partial specification of this referent is a non-human thing that can be XEQTed. ("XEQT" is an abbreviation for "execute", which is the action of starting a computer program). Given that XEQT was known to be an action applicable to programs, and that RUNOFF is known to be a program, this semantic approach could find RUNOFF as the referent.

There is a slight problem with the "it" at 1.13. For the correct referent to be found, the system would have to regard RUNOFF as an animate agent, since the referent is specified to be something which "grabs".

For the first Expression in turn 2, the semantic approach has the same problem as discussed for the syntactic approach here: it may find "no output" in turn 1 as co-referential.

The second Expression "the line printer" isn't a repeated reference, but the semantic approach may handle this correctly, depending on the exact comprehension and reference processes it has. If the knowledge by both participants about RUNOFF and TENEX systems includes line printers, and if this knowledge is accessed in understanding the dialogue to this point, then the correct referent for this Expression will be correctly found, even though it hadn't been mentioned previously.

In turn 3, the semantic system may see "the line printer" as co-referential to "it" on the basis of pure recency. The semantic system may be able to reject this, depending on what it knows about line printers and files. But this is a case in which syntactic clues would help a pure semantic system, since these clues would suggest "the output" as a better possibility.

There are cases in which ignoring syntax will cause a pure semantic system to fail.
"... there is a system that you can use that will let you manipulate your unsent mail. *It* is called mailstat, I believe. ..."
In this case, there are least two possible co-referential expressions to "it", "a system" and "your unsent mail", both of which fit the constraints. On syntactic grounds, we can see that "a system" is the co-referential expression. Clearly, we want to combine all the available syntactic, semantic, and pragmatic cues to repeated reference into one integrated approach.

3. Classification of the Bag of Repeated Reference Tricks

Let us look at a set of repeated reference tricks, shown in figure 3. These include the ones we have discussed already, plus additional repeated reference heuristics described in papers by Warnock (1972) and Baranofsky (1970).

We can classify the bag of repeated reference tricks into two categories, aspects of the possible Co-referential Expressions and aspects of the current Expression.

1. Proximity: The closer an utterance containing another Expression is to the current Expression, the more likely the other is to be Co-referential with the current one.
"I can't find any documentation on the program. I have a tape here at Purdue and I can't figure out what format *it*'s in."
The "it" refers to the tape, not to the documentation or to the program.

2. Syntactic role in the sentence:
2a. Subject/Object/Preposition phrases: The syntactic subject of an utterance is more likely to be an Co-referential Expression than the syntactic object, which is more likely than the preposition phrases.
"O: The output comes out on the line printer.
L: Throw *it* away ... "
The "it" refers to the output (the syntactic subject of the first utterance) rather than to the line printer (the syntactic object).

2b: Superordinate/subordinate: Concepts expressed in a superordinate clause are more likely to be Co-referential that those in any subordinate clauses.
" ... the tape that file is archived on seems to be a bad tape. We can't seem to get *it* to read ... "
The "it" refers to the tape rather than to the file (in the subordinate relative clause).

2c: Topicalization: Some special syntactic constructions (such as cleft sentences) can be used to emphasize one element of an utterance. These emphasized Expressions are more likely to be co-referential than unemphasized ones.
" ... there is a background job running here that checks to see if there is any unsent mail. Once *it* finds some, *it* tries to resend it."
The two "it"s are co-referential with the topicalized "background job" rather than with the "unsent mail".

3. Centrality: an concept which has previously been referenced more than once is more likely to be referenced again than one referenced only once.
"L: ... Any chance I can recover [file name] from the most recent system dump?
O: Probably, let me look for *it* and get back to you, ok?
L: Could you SNDMSG to me, one way or the other? I won't be doing anything about *it* tonight. If *it* is there, I will be forever grateful to recover *it*."
The "it"s by L all refer back to L's file, rather than to the system dump or the one way or the other, at least partially due to the previous reference by O.

4. Current topic: An Expression which refers to a concept in the current topic is more likely to be Co-referential. Deutsch (1974) observed that repeated reference can normally be made only to concepts that are part of a currently open topic. Once a topic is closed, it must be reopened before concepts within it can be referenced again.

"L: ... Can you recover those files for me..as far as I know they were in the directory on the 16th...the names are ...

[intervening dialogue]

O: OK I have found ***the files you want*** I will retrieve ***those*** for you ... "

In the second utterance, O had to initially specify the files in some detail, but once the topic was re-established, she could use just "those".

Figure 3a: Aspects of Possible Co-referential Expressions

1. Specification by the current Expression:
   1a: The the pronoun or determiner in an Expression often specifies that the referent concept be a particular number and/or gender.
   "How do *I* get Runoff to work?"
   The pronoun "I" completely specifies the referent to be the speaker.

   1b: The other words in the current Expression often further specify the referent.
   "I have found *the files you are concerned about ...* "

2. Specification by the verb in the same utterance: The verb in the same utterance as the Expression often specifies that the referent have certain properties.
   "Throw *it* away."
   The verb (and verb particle) "throw away" specifies the referent to be something of little value to the speaker.

3. Specification by the whole clause containing the current expression: Sometimes the particular combination of a verb and its noun phrase arguments puts constraints on the referent.
   "Did *it* produce any output file?"
   The combination of the verb and the object limit the referent to being a computer program of some kind.

Figure 3b: Aspects of the Current Expression

3.1 The Possible Co-referential Expressions

The heuristics that deal with aspects of the possible co-referential expressions are a diverse lot, covering both syntax and semantics. One thing seems to be a common feature - each seems to reflect contributions to the salience of the various concepts in awareness. Some of the heuristics capture what is called "focus"; others reflect the fact that concepts in awareness are temporary, disappearing if they aren't repeatedly refreshed.

3.2 Aspects of the Current Expression

Initially, the referent of an Expression is completely unspecified . Each of the repeated reference heuristics given in the figure 3b can be seen as contributing specifications to this unspecified concept. Each heuristic may operate independently in adding its constraints to the referent of the current expression. And the end result is the partially specified referent.

This classification of the repeated reference heuristics into these two categories is straightforward. However, it suggests a general approach toward modeling repeated reference abilities in process models. There are two parts of a model - all those processes that contribute specifications to the referent of the current expression, and the processes that affect the salience of all the other currently active concepts.

4. An Activation Model of Repeated Reference

Let us assume we have some standard parser, that takes an utterance and chunks it up into Expressions.

Let's put each of these units into one place, called a Workspace, and give it a numeric value, called its Activation Rate. This Activation Rate reflects the momentary salience of the unit (it is similar to the "importance" metric discussed by Warnock (1972)). When a comprehension process attempts to put an Activation in the Workspace (to represent some new intermediate result), the Activation Model will first look for an identical existing activation. If none exists, the Model will create a new Activation with a specified Activation Rate. However, if an existing Activation is found, the Model instead increments its Activation Rate by the specified amount.

There is a Threshold value for existence of an Activation. Those Activations with rates below the Threshold are removed from the Workspace.

One part of the Activation Model is a set of rules for modifying the Activation Rates, derived from the heuristics in figure 3a. These rules are given below. The contents of the Workspace will be undergoing continual change, with new activations being created, and existing activations being deleted whenever they fall below Threshold. The contents at any one moment serves as the current context for reference resolution.

Let us explore the operation of this Activation Model of Repeated Reference by hand-simulating its performance on the dialogue in figure 1. To do this, we first have to specify a number of parameters (in a somewhat ad hoc manner) for the rules given

below. The first part of each rule description covers general aspects, and the second part gives a precise specification (with ad hoc parameters) that will be used to show the operation on the example discourse.

4.1 Rules for Modifying the Activation Rates of Possible referents

1. Recency of utterance rule:
   As each new utterance comes in, decrement the rates of all the preexisting activations.
   For the example below, decrement all existing activations by 50%.

2. Primacy within utterance rule:
   Increment the activation rate of the concept representing each unit of an utterance as it enters, but by a successively smaller amount. A subclaim of this Model is that the three separate syntactic factors 2a (Subject/object/preposition phrase), 2b (Super/subordinate clauses), and 2c (Topicalized constructions) in figure 3a can be captured in this one rule.
   For the example, increment the activation for the first expression by 10 (arbitrary) units, the next by 9, etc.

3. Centrality:
   When a concept has been referenced more than once, the same activation will have been incremented by each reference. Thus it will be more salient than if referenced only once, and so more likely to be selected again as a referent.
   The centrality aspect is already captured by the Activation Model itself, so we need no separate rule.

4. Current topic:
   Given a comprehension process for detecting topic structure, a rule that incremented the activation rates of all components of a newly detected topic would capture the current topic aspect. We don't yet have a model of this process, so we can't use this rule. But this illustrates the way that progress in modeling other aspects of natural language can be easily interfaced to this Activation Model extending its capabilities.
   The operator-linker dialogues, such as the one in figure 1, generally contain only one topic throughout, so the present lack of a topic structuring process isn't very noticeable for reference resolution in the cases we have been considering from these dialogues.

Now what do we do with an Expression? Well, we treat it much like any other unit. We create an activation for it, and start constraining what can fit in this spot by adding specifications. In the particular repeated reference model we are developing here, there are a set of rules for adding to this specification of the Expression, which are derived from the heuristics described in figure 3b.

So far, we have specified the aspects of possible co-referential expressions (figure 3a) as a set of rules for modifying activation rates within the Activation Model. Now we have to specify the other part of the model, capturing the aspects of the current expression (figure 3b).

4.2 Rules for Specifying the Referent

1. Definitions of the reference words:
   The reference words (pronouns & determiners) themselves contribute specification of number and/or gender.
   For the example, we need to define only the following words, so that these specifications can be added to the referent of the expression that they are in.
   "I" or "me" or "my" --> singular and human and current speaker
   "you" or "your" --> human and current hearer
   "it" or "this" --> singular
   "these" or "those" --> plural

2. Explicit modifiers in Expressions:
   If there are other words in the current expression, these further specify the referent. These include adjectives and nouns, prepositional phrases, and subordinate clauses.
   The following are words which are in expressions in the example dialogues and which add specifications to the referent: input, file, output, line printer, commands, comma.

3. The verbs in the same clause as the current expression:
   The kind of action described by the verb often adds further specifications to the referent.
   These are the verbs that occur in the example dialogue, most of which add specifications to their arguments: get, keep, xeqt, grab, give, come out, throw away, confirm, query.

4. The clause that the current expression is in:
   The event described by the rest of the clause also helps specify the referent.

The model then applies a match process between the partial specification of the referent of the current expression and the salient concepts in the Workspace, and selects the most salient concept that best matches the specification.

Let us now examine a hand-simulated operation of this Activation Model on the dialogue in figure 1.

For all the instance of first person pronouns ("I" at 1.3, 1.8, 3.6; "my" at 1.16; "me" at 1.25), the referent will be specified to be the current speaker and the expression will acquire that person as the referent.

The second pronouns ("you" at 4.8; "your" at 4.2) will all acquire the current hearer at turn 4 (the person L) as their referent and thus be co-referential expressions.

Let us focus on the "it"s, since these are the most challenging cases. When the "it" at 1.11 occurs, there are only a few activations in the workspace: person L (rate

4.5) and Runoff (rate 4). "It" specifies the referent to be singular; the verb specifies the referent to be a program. The only match is with Runoff, so that is acquired as the referent concept.

For the "it" at 1.13, the workspace now contains the same activations, but with different rates: person L (rate 7.25) and Runoff (rate 6.5). The specification of the referent are singular, animate, and involved with input files. Again, Runoff is acquired as a referent, making the expressions at 1.3, 1.11 and 1.13 all co-referential. This agrees with the annotations of the observer (figure 2).

Things are a little more complicated for the "it"s in turn 3, since there are more activations in the workspace at that time. These are the activations: specific output (rate 5), specific line printer (rate 4.5), person L (rate 3.3), non-existent output (rate 2), the word "done" (rate 1.1), Runoff (rate 1.9), and specific input file (rate 0.8). The specification of the referent include singular, movable, valueless to person L. The best matches are with the specific input and the specific output, and since the specific output is more salient, it would be acquired as a referent. This differs slightly from the observer's annotation, since he noted "it" as referring to a specific instance of the concept referred to by "the output".

Finally, the referent of "it" at 3.8 will be specified as being singular, movable, and an entity internal to computers. The state of the Workspace will be similar to that for the previous "it", with the Activation rates of the specific output higher and all others lower. The "it" would acquire the specific output as a referent. At one level the two "it"s in turn 3 are co-referential, but their low level referents are different and at this level they are not co-referential. We will examine this issue in more detail in section 7.

5. Complex Repeated Reference

So far, we have been primarily concerned with simple repeated reference, in which the co-referential expressions are simple noun phrases with relatively concrete concepts as referents. However, Expressions are often used to refer in much more complex ways. Our Dialogue Modeling Project at ISI has investigated two kinds of complex repeated reference. Text Reference and Propositional Reference (Mann, et. al., 1975; Archbold, 1975).

People sometimes use Expressions to refer to words or phrases that have been previously said. For example, we can talk about the last sentence or about this sentence - two text references. People more commonly use Expressions to refer to whole propositions that have been uttered previously. For example, I can refer to the previous description of repeated reference tricks - a reference to a concept that spans several pages.

These two kinds of references and their attendant complexities are discussed at length by Archbold (1975). To illustrate some of the many levels of text and propositional reference, consider the following (constructed) examples:

1. Four. That rhymes with score.
2. Four. That is a four letter word.
3. Fourscore and seven years ago. That's eighty seven years ago.
4. Four score and seven years ago. That's the opening phrase of a

famous speech given in 1863.
5. Four score and seven years ago. That was 1776.
6. Four score and seven years ago. That was 1888.

(See Archbold (1975) for a set of real examples that make this same point.)

Basically, people are able to reference a whole spectrum of concepts, including those concepts that represent intermediate results in the comprehension process. Among these are: a particular feature of the pronunciation of words, a specific use of a word, the word in general, a phrase, clause or sentence. These are all clumped under the category of text reference.

At a deeper semantic level, people are able to reference the concepts that represent various levels of comprehension and understanding. These include: actions, simple low level propositions, and larger scale units like topics. These are all types of proposition reference, and, like text reference, the referents are intermediate results of a comprehension process.

5.1 Complex Repeated Reference and the Activation Model

Existing models of repeated reference have had difficulty with these more complex kinds of reference. The few systems that address propositional reference at all handle it in an extremely limited and ad hoc way (for example, Winograd's SHRDLU), and none tackle the problem of text reference.

What about the Activation Model presented here? For this model to find a referent, the concept has to be in the Workspace. If we can get our language comprehension processes to put all their intermediate results into this Workspace, then these results will be temporarily available as possible referents of a text or propositional Expression. The Workspace will contain the whole spectrum of currently active concepts, all of which will fade away if not referenced again soon.

This way of expanding the scope of the activation model fits very naturally within a recently proposed general framework for process models (Levin, 1975). In this framework, called Proteus, all processing, at all levels, takes place within such a Workspace. However, even with more conventional language comprehension models, copies of intermediate results can be added to the Workspace, thus broading the scope of the referential processes.

5.2 Quotes

Quotes generally play a big role in much of the philosophical discussion of text reference. They are generally interpreted as signifying that the word itself is meant, rather than the underlying concept for which the word is the name.

However, in real dialogue, quotes (which occur rarely) are used in a broad variety of ways. (See the discussion of many of these by Archbold (1975).) One reaction to the hodge-podge of actual usage of quotes is to retreat to the performance/competence distinctions of linguists. However, once we take the actual uses seriously, we can see

that quotes are used as a "warning signal" to the hearer that the interpretation of the quoted word or phrase is meant to be different from the usual interpretation. Now, as we have seen in our examination of text and propositional reference, there are concepts at many different levels in the comprehension of a word or phrase that can be referred to. So, as a first approximation, we can extend the Activation Model of Reference to deal with quotes by suppressing the initial interpretation of a quoted word or phrase, thus allowing some less salient concept to be selected as the referent. This hypothesis for modeling quotes thus approximates the wide variety of ways in which quotes are actually used.

6. Non-Repeated Reference

People often use Expressions to introduce concepts not discussed previously. One way in which we do this is to refer to concepts not explicitly mentioned, but which are closely related to those that were. To account for this use of Expressions, Chafe (1972) introduced the notion of "foregrounding", in which the mention of a concept made closely related concepts available for referencing.

A classic (constructed) example of this is: "I rode a train today. I was allowed to toot the whistle". The phrase "the whistle" is a foreground reference. We don't just want to fall back on our default action of using the specification of the referent of the current expression (that this is some whistle that can be tooted). Instead, we are able to further determine that this is a specific whistle, which is controlled from the engine cab, etc. The first sentence "foregrounded" the knowledge about trains, so that the reference to "whistle" in the second can be determined to mean a very particular kind of whistle.

There are two systems which allow kinds of foreground references, Rieger's inference component of the MARGIE system (Rieger, 1974), and the SAM system (Schank, 1975).

In the MARGIE system, Expressions were detected by the parser and passed to the inferencing system for resolution. The set of possible referents considered by the inference system included not only those directly derived from previous utterances, but also those derived from any inference made from these utterances. In this way, the "inferred" concepts were "foregrounded" and thus available for referencing.

The SAM System (Schank, 1975) is an implementation using Scripts (Schank & Abelson, 1975) as a high level organization for language understanding .

In comprehending a particular set of utterances, a Script is found and used to guide further comprehension. Since a Script is an organized body of knowledge, the system can use it to generate expectations of future utterances. More interesting for this discussion, it provides a set of "closely related" concepts, all available for foreground reference.

6.1 Foreground Reference and the Activation Model

The Activation Model of Reference can be extended to deal with foreground reference in the same way as we extended it to deal with complex repeated reference. We put all the intermediate results of our inferencing and comprehension processes into the Workspace. For example, if we use higher level organizing concepts in comprehension, then whenever these concepts are used in comprehending an input, they will be put into the Workspace, and thus all the components will be available for referencing.

This proposed solution illustrates (and utilizes) the dependence of reference processes on other comprehension processes. Whenever some new comprehension process is developed, the capabilities of the referencing processes will also be expanded if the "results" of the new processes are added to the Workspace.

6.2 Failure to Find a Referent

What if we still don't find a referent concept? We already have a partial specification of the referent, and in many cases, this is all we need to know about this concept. Many Expressions seem to need no definite referent at all for comprehension to proceed satisfactorily. These cases are explored by Martin (1975) in some detail. In considering these cases, he developed a generalization of the notion of a simple pointer to a concept as the "referent" of an expression. These are "descriptions" with varying degrees of detail. These descriptions, which are partially specified referents, can be utilized in performing inferences, and also stored as knowledge known about the concept.

7. Referents and Reference

So far in this paper, the term "referent" has been used 128 times. Let us now examine explicitly how this term is being used, and therefore, what position on reference has been implicitly assumed.

We have talked about Expressions as sets of words which have concepts as their referents. One might be tempted to claim that the referent concept for an expression is the "meaning" of that expression. However, we have seen several ways in which this has to be modified. First of all, the particular concept which is the referent of an expression depends on the current context - the same expression can have two different referents in two different contexts. Given this observation, one might be tempted to view the referent of an expression as entirely context-dependent, and therefore that expressions by themselves have "no meaning".

However, our further investigation of complex reference and non-repeated reference have led us to a possible reconciliation of these two views of reference. Instead of a single referent concept, we found that an expression has a whole family of referents at many different levels. As the comprehension of a given expression proceeds, intermediate interpretations of the expression are generated, at successively more abstract levels. The initial referents represent the surface characteristics of the expression - the set of component letters, the shape or sound of the words. Successive levels include the low level semantic specifications - number, gender, the superset concepts.

The lower level referents of a given expression will be the same across context, while the more abstract referents will differ from context to context.

Now, we can consider two expressions to be co-referential whenever they have identical referent concepts at some level. That is, if expressions E1 in context C1 has referents (R11, R12, R13, ... , R1N) and expression E2 in context C2 has referents (R21, R22, R13, ... , R1N), then the two expressions are co-referential because they have the same referent concepts at level three and beyond.

For example, "Marina del Rey" and "*Marina del Ray*" are co-referential a a fairly low level; "La Jolla", "the place where UCSD is located", and "The Jewel of the Pacific" are co-referential at a slightly higher level, and "it" and "Runoff" are co-referential at a higher level in the utterance "How can I get *Runoff* to work? I keep xeqtn *it* ...".

Co-referential expressions are thus "the same" above a certain level, but different below that level. This difference is the reason why statements like "La Jolla is The Jewel of the Pacific" aren't empty tautologies (X is X).

A referent concept is then one of the family of concepts that represent an expression at some level.

8. Summary

In this paper, we have examined a number of the repeated reference heuristics used by language understanding systems. After observing the difficulties that existing models have with repeated reference in real dialogues, we divided these heuristics into two categories, those dealing with aspects of the possible co-referential expressions, and those dealing with aspects of the current expressions. Working from this categorization, we were able to propose an Activation Model for simple repeated reference, and then to extend it to text and propositional repeated reference, and finally to non-repeated reference. Some general issues of reference were examined in light of the Activation Model.

Acknowledgments

I want to thank the other members of our Dialogue Modeling Project at ISI, particularly Jim Moore, who read previous drafts of this paper and helped me clarify the terminology and concepts used here. Some of the ideas in this paper originated in discussion with our summer members, especially Alex Archbold (from Harvard) and Paul Martin (from Stanford).

References

Archbold, A. A., "Text Reference And Repeated Propositional Reference: Concepts And Detection Procedures", ***Working Papers in Dialogue Modeling - Vol. 1,*** USC/Information Sciences Institute, ISI/RR-77-55, January 1977, Sec. 2.

Baranofsky, S. "Some heuristics for automatic detection and resolution of anaphora in discourse", Unpublished M. A. thesis, U. of Texas, 1970

Chafe, W. L. "Discourse structure and human knowledge". In R. O. Freedle & J. B. Carroll (eds.) ***Language comprehension and the acquisition of knowledge,*** Washington, D.C.: Winston, 1972.

Collins, A. M., & Quillian, M. R. "How to make a language user". In E. Tulving & W. Donaldson (eds.) ***Organization of Memory*** New York: Academic Press, 1972.

Deutsch, B. G. "The structure of task oriented dialogs." ***Proceedings of the IEEE Symposium on Speech Recognition.*** Pittsburgh, PA., 1974.

Levin, J. A. "PROTEUS: An activation framework for cognitive process models." Doctoral dissertation, Univ. of Calif., San Diego, 1975.

Mann, W. C., J. A. Moore, J. A. Levin, and J. H. Carlisle, ***Observation Methods For Human Dialogue,*** USC/Information Sciences Institute, ISI/RR-75-33, June 1975.

Martin, P. A. "Why describing is better than pointing." Unpublished paper. Information Sciences Institute, Marina del Rey, CA., October 1975.

Olney, J. C. "An investigation of English discourse structure with particular attention to anaphoric relationships". Santa Monica: System Development Corporation, TM-4229, 1969.

Quillian, M. R. "The Teachable Language Comprehender: A simulation program and theory of language". ***Communication of the ACM***, 1969, *12*, 459-476.

Rieger, C. "Conceptual memory: A theory and computer program for processing the meaning content of natural language utterances". Doctoral dissertation, Stanford Univ., AIM-233, 1974.

Rumelhart, D. E., & Levin, J. A. "A language comprehension system". In D. A. Norman, D. E. Rumelhart, & the LNR Research Group, ***Explorations in cognition,*** San Francisco: W. H. Freeman, 1975.

Schank, R. C. "Using knowledge to understand." ***Theoretical issues in natural language processing,*** Proceedings of an ad hoc conference, Cambridge, Mass., 1975.

Schank, R. C., & Abelson, R. P., "Scripts, plans, and knowledge," Paper presented at the Fourth international joint conference on artificial intelligence, Tbilisi, USSR, August, 1975.

Schank, R. C., Goldman, N., Rieger, C. J., & Riesbeck, C. "MARGIE: Memory, analysis, response generation, and inference on English", *Proceedings of the third international joint conference on artificial intelligence*, Stanford, 1973.

Warnock, E. H. "Anaphoric reference: Finding the antecedent". Unpublished paper, Harvard University, 1972.

Winograd, T. "Understanding natural language". *Cognitive Psychology*, 1972, 3, 1-191.

Working Papers in Dialogue Modeling

Volume 1 - Section 2

Text Reference and Repeated Propositional Reference: Concepts and Detection Procedures

Armar A. Archbold

Original Date: August 15, 1975

REPEATED TEXT AND PROPOSITIONAL REFERENCE: CONCEPTS AND DETECTION
INTRODUCTION

I. INTRODUCTION

When people engage in dialogue, they quite frequently, at some point, refer back to and talk about something that has already been said in the dialogue, or refer forwards to something that is about to be said. In some cases, they may refer to a string of words uttered at some nearby point in the dialogue, thereby making a "*(repeated) text reference*". In other cases, they may talk about some state of affairs, some statement, some belief, which is described, made, expressed or referred to nearby in the dialogue, thereby making a "*(repeated) propositional reference*".

Both text reference and repeated propositional reference are of interest to the Dialogue Analysis Project. Two questions immediately arise concerning these dialogue phenomena, however. First of all, there is the conceptual problem: (a) is it possible to define these phenomena at all clearly? (b) if there are several different feasible definitions available, which should the Dialogue Analysis team choose, given its goals and interests? Secondly, there is the detection problem. Given a definition of these two phenomena, how can one determine what expressions may be involved in them, and which particular expressions are involved in a particular instance of them?

The present paper addresses these two problems. The first section will survey some philosophical and linguistic literature's treatment of the notions of "text" and "proposition" and of problems associated with these notions. The available notions will then be discussed and evaluated in the light of the team's interests and goals, and in view of some of the data that the team will have to account for. The second section will discuss various "clues" which might be relied upon to determine whether a given expression is being used in a dialogue to make a text reference or a repeated propositional reference. The focus will be upon the question: are there any operationalizable procedures for detecting pairs of expressions which are involved in repeated text or propositional reference?

REPEATED TEXT AND PROPOSITIONAL REFERENCE: CONCEPTS AND DETECTION
THE NOTIONS OF TEXT AND REPEATED PROPOSITIONAL REFERENCE

SECTION I:

CONCEPTS OF TEXT REFERENCE AND REPEATED PROPOSITIONAL REFERENCE

In this section, the concepts of text reference and of repeated propositional reference will be examined, first as they are presented in some philosophical and linguistic traditions, and then in view of their operationalizability for the team.

Below, we shall use the following terminology and abbreviations:

REXPR	a referring expression, i.e. an expression which is used by a speaker in a dialogue to refer to, mention, or pick out an "object" or a set of "objects" (where "object" is taken in its widest possible sense, to include physical objects, people, states of affairs, events, actions, processes, abstract constructs such as the numbers or the quality of beauty, intentional objects such as thoughts or beliefs, hypothetical or fictional objects such as the child X and Y would have conceived together had they not separated, or Pegasus, and so on).
TREF	text reference
PREF	propositional reference
RREF	repeated reference: the use of two separate linguistic expressions to refer to (designate, mention, pick out, etc.) the same "object". The first referring expression involved we call "the antecedent referring expression" (AREXPR); the later of the two expressions involved we call "the consequent referring expression" (CREXPR).
RPREF	repeated propositional reference.

I.1 TEXT REFERENCE

I.1.1 NOTIONS OF TEXT REFERENCE IN THE LITERATURE

In the literature, the notion of text reference is closely bound to the distinction between the use of language and the mentioning of language. We ordinarily *use* language to talk about the world ; but we may also use language to talk about language, to discuss linguistic expressions - words, phrases, sentences, or whole texts or discourses. In the latter case, we ***mention*** or refer to linguistic expressions in order to say something about their phonetic, orthographic, syntactic, semantic, logical or pragmatic features. [1]

REPEATED TEXT AND PROPOSITIONAL REFERENCE: CONCEPTS AND DETECTION

THE NOTIONS OF TEXT AND REPEATED PROPOSITIONAL REFERENCE

Traditionally it was thought that the use and mention of expressions were very distinct operations, and that to ignore the distinction was to run the risk of creating such nonsense as:

My dog contains three letters.

or

My "dog" is a hearty eater.

Though it was allowed that one could mention expressions by the use of proper names or definite descriptions, the traditional paradigm of text reference was text reference by quotation. A quotation-expression - i.e. a quoted expression plus its surrounding quote marks - is usually analysed as a name for the expression within the quote marks.

There is an important ambiguity in the notion of mentioning expressions. Text reference is not reference only to individual sounds or inscriptions. When we refer to linguistic expressions, we may refer either to tokens or to types. *Tokens* are particular, unique strings of marks or sounds. *Types* are not single homogeneous entities (there is no such thing as a type-word), but are rather classes of individual tokens, grouped together by reference to some set of taxonomic criteria [2], which can be referred to by the use of

[1] The fact that natural language can be used to talk about natural language expressions (to "talk about itself" in some sense) leads to well known antinomies. Consider the next sentence. The third sentence of footnote 1 of this paper is false. Is the preceding sentence true or false or both or neither? In order to avoid such paradoxes, logicians who have talked about linguistic expressions in order to define their truth conditions have distinguished between the language they examine - the object-language - and the language they use in their analyses - the meta-language. Cf. Alfred Tarski, "The Semantic Conception of Truth", in Leonard Linsky, ed., "Semantics and the Philosophy of Language", University of Illinois Press, Chicago, 1952.

[2] Two remarks are in order here. First, we may very well be unable to explicitly specify some or most of the criteria which we use to group tokens into classes, i.e., to decide when two different tokens are of the same type. Secondly, it is probably wrong to thing that we sort tokens into types by reference only to orthographic or phonic criteria, without any consideration of their semantic role. What we refer to when we perform a text reference to a token is a sound or a mark which belongs to a language, which was produced for a characteristic linguistic purpose, or was produced in a context in which describing the sound or mark in terms of its syntax, relative to a framework in which such events or marks can be systematically described, is an appropriate activity to engage in. And we disregard orthographic differences between tokens which would have no impact upon their semantic roles... Cf. D.V.C. Lincicome, "Systematically Ignored Differences and the Identity of Propositions", Foundations of Language, Vol 12, No. 1, September 1974, Section 2.

species-words similar to "dogs" (which refers to the class of all dogs). A convention which one might adopt to avoid confusion here is to refer to tokens by singling them out by means of ostension or definite description, and then christening them with a proper name. For example, the mark on the next line

animal

one could call Tom. Species-words can then be used to refer to classes of such individual tokens; species-words could be formed by spelling the token backwards, inserting dots between the letters.
Thus if the two marks on the next line

animal towards

are christened Dick and Harry, one could say that both Tom and Dick are l.a.m.i.n.as - that is, they are both tokens which are of the l.a.m.i.n.a class or species [3]. Given the distinction between tokens and types, one must clarify one's analysis of the role of quotation-marks, and separate out those cases in which quotation-marks form the name of a token, and when they form the name of a type to which the individual quoted expression token belongs.

Let us consider some examples of TREF. The examples below - all partly or wholly concocted for purposes of illustration - involve two or more underlined expressions which refer to the same text; such repeated reference is simply designed to highlight the TREFs being performed.

(1) A: The third letter from the top of the list on the blackboard is "D".
B: Yes, I know. It's written in red chalk. But I can't make out the letter directly below it.

(2) A: John Smith is calling himself "Hiroto Texagewish" these days.
B: His assumed name is certainly hard to pronounce!

(3) A: On checklist G/8-2 under step 5, there's a statement "disable all jets on two adjacent quads". Is that what you are talking about?
B: No, "disable all jets on two adjacent quads" is not what I was referring to.

(4) A: SHOULD I TYPE ATT LINKER (PASSWORD), JOB NUMBER?
B: YES, THAT'S WHAT YOU SHOULD TYPE IF YOU WANT TO RE-ATTACH.

(5) A: John gave me a good example of a tongue-twister yesterday. What was it now? Ah yes, I remember. How much wood would a wood-chuck chuck if a wood-chuck would chuck wood?.
B: Well, the tongue-twister he proposed was easier to pronounce than the one Mary came up with.

[3] L. Goddard and R. Routley, "Use, Mention and Quotation", The Australasian Journal of Philosophy, Vol. 44, No. 1, May 1966.

(6) A: You're coming in a little louder now.
B: Fred, you'll have to say that again. I couldn't hear you; there's too much background noise.
A: Okay. I said, you're coming in a little louder now.

Consideration of the above examples leads to the following remarks:

(1) The linguistic expressions which are used to refer to (mention) other linguistic expressions are of two types. The first type *exhibits* or *replicates* the expression to which it refers. Thus in the two following sentences

John wrote "Today's lesson is Chapter V" on the board.
Should I type TTY or TTY:?

the underlined expressions are used to refer to expressions which they replicate or exhibit. The difference between exhibiting and replicating can be shown by the following examples:

The following letter, "D", was not written by hand.
(exhibition)
Yesterday, John wrote "Today's lesson is Chapter V" on the blackboard.
(replication)

In the first case, one is exhibiting a token and talking about that very token, and no other. In the second case, one is talking about a token John produced yesterday, and one is doing so by exhibiting a token of the same type as the token he produced yesterday; in that sense one is replicating a token, by producing and exhibiting a "copy", a token of the same type.

The second type of linguistic expression which is used for TREF neither exhibits nor replicates the expression to which it refers. An example of such an expression would be "The sentence John wrote on the board yesterday".

(2) Exhibiting or replicating text referential expressions may or may not use quotation marks: compare the examples (1-3) and (4-6). This fact flies in the face of the traditional view according to which one only exhibits or replicates expressions to which one wants to refer by using quotation-marks, quotation-marks being seen as operators which form a proper name of the string enclosed within them. One can exhibit or replicate a linguistic expression not only by quotation, but also by capitalization, pauses, indentation or spacing, intonation and many other ways besides. There thus seems to be no one reliable orthographic sign of exhibition or replication; there may, however, be a list of such signs. In any case, the function traditionally assigned to quotation marks can be performed by many other marks or sounds. (One should also note that quotation marks are sometimes used in complex ways which bear only a faint resemblance to their stereotypical use, as when they are employed in ironical remarks, as shudder quotes or snigger quotes. Examples: "These days, one never gets price stability, only decreases in the rate of increase of inflation. That's "progress".", or Lenin's statement "We will "support" the Mencheviks as the rope supports the neck of the hanged man!")

(3) Text referential expressions which are not of the "exhibition" or "replication" types may be (a) pronouns or (b) noun phrases. The noun phrases may be proper nouns (think of referring to a string of marks on a piece of paper as "Tom"), (modified) common nouns ("a password", "the loud shout") with or without relative clauses.

(4) Text referential expressions which are not of the "exhibition" or "replication" type may be combined with those that are to form complicated text referring expressions. Consider example (4), where "(PASSWORD), job number" are used alongside two replication text expressions. It is interesting to note that the order in which these referring expressions occur is the same as the order in which the referents of each individual expression must occur in order to form a token of the type which the entire string "ATT LINKER (PASSWORD), job number" denotes.

One might conclude that text reference is a *pure* mention of phonetic or orthographic tokens or classes of tokens. However, the traditional clearcut distinction between use and mention has been attacked of late - and justifiably so. Things are not so simple.

> "There are ... many sentences in which an expression is both introduced and is also used; in particular, sentences which are used to convey both linguistic and factual information. This is especially so in sentences containing the words 'call', 'distinguish', 'determine', 'is called', ... 'satisfy', or compounds of such words. Consider, e.g. 'That sleek red-coated dog is Rover', 'The "Queen Elizabeth", which is so-named (so-called) after the present Queen Mother, sailed for Southampton yesterday', 'What is halva?', 'Call her a shrew',
>
> 'If triangles are taken as three-sided figures then they have...', and also indirect speech forms in which the speakers actual words are reported."[4]

There are sentences containing text references in which the text refered to must not only be considered as an uninterpreted orthographic object, but also be "read with understanding", i.e. interpreted. Some examples of such sentences are:

(1) The sign says, "George Washington slept here", but
I don't believe he ever did.
(2) Whenever Fred sighs "Boy, do I need a drink", he expects
you to fix him one.
(3) What he actually said was "It's clear that you've given
this problem a great deal of thought", but he meant
quite the opposite.
(4) "I talk better English than both of youse", shouted
Charles, thereby convincing me that he didn't.

In all of the above, the presence of pronominalization, ellipsis or semantic anaphora

[4] L. Goddard and R. Routley, op.cit., pg. 22.

involving terms both within and outside of the quotation marks shows that the quoted sentences must have been interpreted.[5] Thus if we say that text reference involves the mention of a linguistic expression, *we must allow that such mention may be accompanied by use, and that therefore we must count as text reference not only reference to tokens or types alone, but also reference to tokens or types paired with their meaning, some interpretation, the statement they were used to make on some occasion, or the proposition they express.*

I.1.2 THE TEAM'S INTEREST IN TEXT REFERENCE.

In light of the concepts introduced above, let us now consider the notion of text reference which the team is interested in.

In "Observation Methods for Human Dialogue", we find the following explanations of the concept:

> "...a so-called 'Text Reference', made to a string of words in the preceding dialogue itself (and not to the referent of that preceding string of words!)." (pg. 22)
>
> "A Text Reference occurs whenever reference is made to previously occurring words within the transcript. For example, in the sentence 'Go 3 blocks and turn North; by North I mean towards the mountains.' the second use [i.e. occurrence - A.A.] of "North" is a Text Reference to the first. We call this a Text Reference because it refers to the previous use [occurence] of the word itself, rather than to its meaning." (pgs 26-27)

These passages indicate that the team is interested only in singular text references to words previously used in the dialogue. Note that if we interpret these indications strictly, we must conclude that only TREF to previously used tokens are to be focused upon - since word-tokens, and not classes of word-tokens, are used at a particular time and place by participants in a dialogue. On this strict interpretation, the example given, viz.

Go 3 blocks and turn North;
by North I mean towards the mountains.

is of interest only if one reads the second sentence as equivalent to 'I intended to use the token of type "North" which occured in my previous sentence to mean towards the mountains.' It would not be of interest if the second sentence were read as equivalent to 'All tokens of type "North" which I utter I use to mean towards the mountains', for in that case the second occurence of "North" would refer not to the previous token but to a class of tokens of which the previous token is a member.

However, examination of the examples given in "Observation Methods in Human Dialogue" and discussion with team members indicate that they are concerned with a much wider variety of phenomena. What they are concerned with and what they wish to refer

[5] Barbara Hall Partee, "The Syntax and Semantics of Quotation", in S.R.Anderson and Paul Kiparsky, "A Festschrift for Morris Halle", Holt, Rinehart and Winston, New York, 1973.

to by the term *"Text Reference" consists of:*

(1) Singular text reference where an expression is used to refer to either (a) a token which is used either before or after it in the dialogue, or (b) a class of tokens of which one or more tokens used in the dialogue are members.

(2) Repeated text reference where the AREXPR and the CREXPR refer to either (a) the same token, or (b) the same type, or (c) a type and a token which is a member of that type, or, lastly, (d) two token-classes (i.e. types), one of which is a subset of the other. Either or both of the referents of the AREXPR and the CREXPR may either be or contain as a member a token used in the dialogue.

We may repeat here what we have stressed above: if an expression E1 is used in a dialogue, and is elsewhere referred to by means of another expression E2, E2 is a text-referring expression which is involved in a singular text reference (in the sense of (1)) if it is used to refer to E1 as a token, i.e. if it is used to refer to some inscriptional or orthographic features of E1. However, E2 may also (at the same time) be used to refer to the meaning, the referent, or some other non-textual feature of E1. Text reference does not exclude interpretation.

REPEATED TEXT AND PROPOSITIONAL REFERENCE: CONCEPTS AND DETECTION
THE NOTIONS OF TEXT AND REPEATED PROPOSITIONAL REFERENCE

I.2.1 PROPOSITIONS AND PROPOSITIONAL REFERENCE

Whereas the notion of "text" is relatively unproblematical and unconfusing, given certain simple distinctions, the notion of "proposition" is a difficult one to formulate clearly. Unlike "text", "proposition" is very much a technical philosophical term. We find three main concepts of what a proposition is, formulated and used by (1) linguists, (2) logicians or philosophers primarily concerned with logical matters, and (3) by speech-act theorists.

1.2.1.1 Those linguists who make use of the term "proposition" often equate it with the ***meaning, reading or semantic interpretation of a sentence*** . Let us consider for example the Katzian tradition of transformational semantics. In this tradition, the meaning or semantic interpretation of a sentence is a set of sets of structured markers which are assigned to the sentence on the basis of (a) the semantic markers assigned to the component words by a dictionary, (b) the syntactic structure of the sentence, and (c) a set of semantic "projection rules". The semantic interpretation of a sentence is a theoretical construct which is adequate if in conjunction with the rules of a semantic theory can predict the semantic properties of sentences (such as synonymy, ambiguity, redundancy, presupposition, entailment, and so on). Each set of structured markings is a "reading" or a "proposition". If a sentence is assigned several sets of structured markings it is ambiguous and is said (according to the semantic theory which assigns the markings) to ***express several different propositions*** .

> "Sentences are frequently ambiguous, that is, they express more than one sense. Thus, we shall frequently say that an n-way ambiguous sentence expresses n distinct propositions. (We have taken the term 'meaning' to refer to the sum of the propositions expressed by a sentence ...). We also understand 'proposition' to convey what synonymous sentences have in common by virtue of which they are synonymous. Sentences that are synonymous on a sense [i.e. on one of their readings - A.A.] are thus said to express the same proposition, and fully synonymous sentences are said to express the same set of propositions. Semantically anomalous sentences express no proposition at all."[6]

Note that if one regards propositions as the 'readings' of sentences, one allows that not only declarative, but also interrogative, imperative and hortatory sentences express propositions.

1.2.1.2 The logicians' notion of proposition is distinct from the linguistic notion, at least prima facie. Logicians are concerned with formal constraints on inference, and with the notions of truth and falsity which are needed to account for formal validity of arguments. They have thus made use of a concept of "proposition" which is tantamount to the notion of a ***"truth-vehicle"*** :

> "On the logical account of propositions, propositions have been regarded, first and foremost, as truth-vehicles. That is, propositions are taken to be either the things or some of the things which are true or false. There are several

[6] Jerrold J. Katz, "Semantic Theory", Harper and Row, New York, 1972, pg. 120.

> motivations for this view of propositions. One historically important motive has been to provide a subject matter for logic, something for logic to be about. Logic is, in the first instance, the study of inferences. Whether an inference is valid or invalid depends neither on the particular subject of discourse nor on the determinate mode of linguistic expression but solely upon the formal relations between premises and conclusion. Propositions may thus be conceived as sorts of entities which stand necessarily in such relations as entailment and contradiction, and it is these relations which constitute the grounds of valid and invalid inference and are reflected in particular linguistic embodiments. ...
>
> A second motivation for the logical account of propositions is found in the classical correspondence theory of truth. On this view, truth is regarded as a relation between what is the case in the world, the facts, and the thing, whatever it is, which is true. Propositions have traditionally been cast in the role of the second term of this relation." [7]

Thus on the logicians' view, as on the linguists', there is a distinction between sentences - i.e. strings of inscriptions or sounds which belong to some language - and the propositions which sentences express: Thus the following distinct sentences (taken either as tokens or types)

(1) The moon is smaller than the sun.
(2) The sun is larger than the moon.
(3) La lune est plus petite que le soleil.
(4) Le soleil est plus grand que la lune.

would be said to express the same proposition, and to be true because they all express the same true proposition.

One may to some extent distinguish between a traditional logician's notion of a proposition, and more recent notions.

1.2.1.2.1 My own stereotypical characterisation of the traditional notion is as follows:

(1) a proposition is an abstract object;

(2) there are non-denumerably many propositions;

(3) propositions exist independently of language (ie. there are many propositions which are not, may never be, and perhaps could not be, expressed by a sentence;

(4) a proposition is something which is itself true or false in an absolute (timeless) sense;

(5) some propositions are expressed by declarative sentences;

[7] Rosenberg, Jay F., and Travis, eds., Charles, "Readings in the Philosophy of Language", Prentice-Hall, New Jersey, 1971, pgs.219-220.

(6) if two sentences express the same proposition, then they are analytically equivalent (ie. their meanings constrain their truth conditions in such a way that they are either both true or both false in any situation, state of affairs, or "possible world");

(7) however, if two sentences are analytically equivalent, they need not express the same proposition: analytical equivalence is less stringent a requirement than identity of proposition expressed (thus though 'John is a bachelor' and 'John is an unmarried adult male human being' are analytically equivalent, we do not want to have to say that they express the same proposition);

(8) if two sentences are logically equivalent, i.e. have the same truth-values in all models, then they need not express the same proposition (for example, '$x=x$' and 'Fx or $\sim Fx$' are logically equivalent, because true in all possible worlds, but we do not want to say that they express the same proposition);

(9) propositions are what are believed, doubted, hoped for, etc., i.e. they are the objects of belief, doubt, hope and the other so-called "propositional attitudes".

1.2.1.2.2 The traditional logical notion of proposition, though it can be made precise to some degree [8], has been much criticised as ontologically unnecessary and obscure [9]. Logicians are interested in what is true or false. What we usually term true or false are ***sentences uttered by people in certain contexts and interpreted in certain ways***. It was argued that to postulate the existence of propositions above and beyond uttered sentences was simply to complicate furthur an already vexing question with such additional quandries as the exact nature of the relationship between propositions and the sentences which "express" them. It was felt that the real problem is to determine the nature of the dependency of the truth of sentences upon the context of their use:

> "A sentence is not an event of utterance, but a universal: a repeatable sound pattern, or repeatedly approximable norm. Truth cannot on the whole be viewed as a trait, even a passing trait, of a sentence merely; it is a passing trait of a sentence for a man. 'The door is open' is true for a man when a door is so situated that he would take it as the natural momentary reference of 'the door' and it is (whether he knows it or not) open. The individual event of utterance can still be described as true absolutely, since a time and a man are specific to it; but talk of sentences as true for men at times covers more ground, for it includes cases where the sentence is not uttered by the man in question at the time in question.

[8] Cf. Jan Berg, "What is a Proposition?", Logique et Analyse, Vol. 10, Dec. 1967 (summarized).

[9] Cf. the arguments summarized in Howard Pospesel, "The Non-Existence of Propositions", The Monist, Vol. 53, April 1969 (summarized).

> Relativity to times and persons can be awkward on account of the supplementary specifications in which it keeps involving us. This is no doubt one reason why philosophers have liked to posit supplementary abstract entities - ***propositions*** - as surrogate truth-vehicles."[10]

If a person utters a sentence in a certain manner in a certain place, at a certain time, to a certain interlocutor, in brief, in a certain context and "cotext" (verbal or dialogue context), he has said something, ***made a statement*** , by ***using a sentence in a certain way*** [11]. Many contemporary philosophers regard statements as "primary truth-bearers"; they regard statements as what are (timelessly) true or false. Some further claim that statements are representable by ***eternal sentences*** , i.e. sentences the values of all of whose indexical terms have been explicitly specified, whose truth-value consequently stays fixed through time and from context to context. One might thus think of a statement as a pair comprising a sentence and a complete interpretation of that sentence; the interpretation of the sentence would ideally provide a complete function from possible states of affairs (possible worlds) to truth-values for that statement. Such a function, in other words, would be a complete and precise specification of the truth-conditions of that sentence as used; it would include a set of specifications of truth-conditions which stem from the conventional semantic meaning of the sentence taken just as a sentence of the English (or other) language it belongs to, plus complete specifications of the values of all the indexical terms in the sentence, given the context of use, and unambiguous definite descriptions of the referents of the ambiguous or vague definite descriptions in the sentence. Such interpretations correspond to what some formal logicians have called intensions in their models. Such complete specification is possible by fiat in the domain of formal semantics. It is a moot point whether such a complete specification is possible for a natural language sentence uttered in everyday circumstances.

It is important to note that philosophers may speak of propositions being expressed by sentences, and of statements being made by uttering sentences, but they do not talk of sentences denoting or referring to propositions - except in one case. They do talk of nominalized sentences in modal or intentional contexts as being "proposition-denoting expressions". Consider the sentences "It is impossible that Mary is sick" and "John believes that Mary is sick". In both of these cases one finds the nominalized sentence "that Mary is sick". A traditional analysis of the logic of such sentences claims that the expression "that Mary is sick" is an expression which denotes the proposition expressed (but not denoted) by the sentence "Mary is sick" when occuring outside such "opaque" contexts as modal or (especially) intentional contexts.

[10] Willard V.O. Quine, "Word and Object", The M.I.T. Press, 1960, pgs. 191-192.

[11] For important articles which make use of this terminology, cf. E.J. Lemmon, "Sentences, Statements and Propositions", in J.F. Rosenberg and Charles Travis, eds., 'Readings in the Philosophy of Language', Prentice-Hall, Englewood Cliffs, New Jersey, 1971, and the two articles by P.F.Strawson - "On Referring", Mind, 1950, and "Identifying Reference and Truth-Values", Theoria, Vol. XXX, 1964. All of the three above-mentioned articles are summarized.

THE NOTIONS OF TEXT AND REPEATED PROPOSITIONAL REFERENCE

The two differences between the linguistic and the logical notion of 'proposition' (or 'statement') which are most relevant to us are well summarized by Katz:

> "...[a] point about our [i.e. some linguists'] use of the term 'proposition'... is that the class of propositions cannot be identified with the class of statements, where statements are understood as the logical objects that are the bearers of truth values (i.e. as the objects that obey the law of the excluded middle). The fact that our characterization of the class of propositions encompasses a multitude of nonassertive propositions (questions, requests, etc.), for which it makes no sense to talk about truth and falsity, makes this amply clear. But we cannot even identify the class of statements with the class of assertive propositions, since a proposition with a token indexical element cannot have a fixed truth value."[Ibid., pgs. 122-123]

The third notion of 'proposition' which one can pick out of the literature is that of the speech act theorists. To put it rather vaguely, for speech-act theorists, the proposition expressed by someone who utters a sentence-token is ***what is left, syntactically and semantically, in the uttered sentence token after all of its constituents relevant to a determination of its illocutionary force have been abstracted from it*** [12]. Let us examine Searle's notion of proposition which is of this type.

> "Imagine a speaker and a hearer and suppose that in appropriate circumstances the speaker utters one of the following sentences:
>
> 1. Sam smokes habitually.
> 2. Does Sam smoke habitually?
> 3. Sam, smoke habitually!
> 4. Would that Sam smoked habitually.
>
> ...anyone who utters one of these can be said to have uttered a sentence formed of words in the English language. But clearly this is only the beginning of a description, for the speaker in uttering one of these is characteristically saying something and not merely mouthing words. In uttering 1 a speaker is making (what philosophers call) an assertion, in 2 asking a question, in 3 giving an order, and in 4 (a somewhat archaic form) expressing a wish or desire. And in the performance of each of these four different acts the speaker performs certain other acts which are common to all four: in uttering any of these the speaker ***refers to*** or mentions or designates a certain object Sam, and he predicates the expression "smokes habitually" (or one of its inflections) of the object referred to. Thus we shall say that in

[12] Cf. John R. Searle, "Austin on Locutionary and Illocutionary Acts", in J.F. Rosenberg and C.Travis, eds., op.cit., E.Stenius, "Mood and Language Game", Synthese, Vol.17, 1967, Lennart Aqvist, "Semantic and Pragmatic Characterizability of Linguistic Usage", Synthese, Vol.17, 1967.

> the utterance of all four the reference and predication are the same, though in each case the same reference and predication occur as part of a complete speech act which is different from any of the other three. We thus detach the notions of referring and predicating from the notions of such complete speech acts as asserting, questioning, commanding, etc., and the justification for this separation lies in the fact that the same reference and predication can occur in the performance of different complete speech acts. Austin baptized these complete speech acts with the name "illocutionary acts"...
>
> Whenever two illocutionary acts contain the same reference and predication, provided that the meaning of the referring expression is the same, I shall say the same proposition is expressed. Thus, in the utterances 1-5, the same proposition is expressed. And similarly in the utterances of:
>
> 6. If Sam smokes habitually, he will not live long.
> 7. The proposition that Sam smokes habitually is uninteresting.
>
> the same proposition is expressed as in 1-5, though in both 6 and 7 the proposition occurs as part of another proposition. Thus ***a proposition is to be sharply distinguished from an assertion or statement of it*** , since in utterances of 1-7 the same proposition occurs, but only in 1 and 5 is it asserted. Stating and asserting are acts, but propositions are not acts. A proposition is what is asserted in the act of asserting, what is stated in the act of stating. The same point in a different way: an assertion is a (very special kind of) commitment to the truth of a proposition. ...
>
> I might summarize this part of my set of distinctions by saying that I am distinguishing between the illocutionary act and the propositional content of the illocutionary act." [13]

Searle proposes an analysis of uttered sentence tokens which would distinguish between (a) an illocutionary force indicator, representing those aspects of the uttered sentence relevant to the determination of its illocutionary force (such as the presence of certain performative verbs, word order, stress, the mood of the verb, and so on), and (b) a propositional indicator, representing that aspect of the utterance which is neutral to illocutionary force, viz., the proposition expressed. This propositional indicator might also be called a "sentence radical".

We have a wealth of different notions of 'proposition' to choose between. Rather than debate their respective values, we must now ask what notion is closest to that which the team would like to investigate.

[13] John R. Searle, "Speech Acts", Cambridge University Press, 1969, pgs. 22-30.

I.2.2 THE TEAM'S NOTION OF (REPEATED) PROPOSITIONAL REFERENCE

The dialogue analysis team is at present seeking to formulate a notion of proposition congruent with its research interests, and to employ that notion in selecting a set of phenomena which it feels it is presently ready to examine.

I.2.2.1 The team's notion of proposition. No description exists in print as yet of what the team's concept of proposition or of propositional reference might be (propositional, as opposed to text reference, was not mentioned in the "Observation Methods" report). The following remarks are therefore based on discussions I have had with members of the team, particularly with Jim Levin.

It would seem that the team is moving towards a notion of proposition which is much more akin to that of Searle and of some linguists than to that of the logicians'. The team is employing a notion of proposition as a theoretical notion employed in the context of the modeling of dialogue by means of semantic nets. This notion is to be understood by reference to a certain form of representation of utterances in a dialogue. In order to see this more clearly, consider the utterances in (1)-(3) below, accompanied by one form of representation which the team might employ. (In these examples, we suppose that Bill is addressing his utterances to John.)

(1) Bill: You will shut the door.

```
(SAY
  (BILL
   TIME-1
   JOHN
   (SHUT
        (JOHN
         DOOR-1
         TIME-2))))
```

(2) Bill: Will you shut the door?

```
(ASK
  (BILL
   TIME-1
   JOHN
   (SHUT
        (JOHN
         DOOR-1
         TIME-2))))
```

(3) Bill: John, shut the door!

```
(ORDER
  (BILL
   TIME-1
   JOHN
   (SHUT
        (JOHN
         DOOR-1
         TIME-2))))
```

The three representations have a common element, namely, (SHUT (JOHN DOOR-1 TIME-2)). This common element is very similar to what Searle calls a proposition: it is a predication abstracted from different illocutionary acts (and not only from statements or assertions). This common element is part of what is a proposition expressed by a declarative sentence, e.g., in the case of (1), the proposition expressed by a sentence such as "Bill asked John at time t1 to shut the door at time t2". This matrix sentence or proposition contains an illocutionary verb ("say", "ask", "order", etc.) and thus conveys not only the proposition expressed by the common element, but also the illocutionary force with which that common element was produced. Notice, however, that the representation is one in which indexicals are filled in, and in which the referents of noun-phrases in the utterances are unambiguously specified (thus DOOR-1 is a GENSYM); in this respect the representation is akin to the logicians' representations of statements.

Given this form of representation, one might define proposition in one or more of several ways. One might reserve the term for the representation of an act of uttering a sentence taken in its entirety. One might term proposition any complete representational unit, i.e. any verb and its arguments; a proposition in this sense would include, of course, propositions in the first sense above. Or one might exclude the matrix representation, and only term its components propositions; these propositions would then correspond to (nominalized) sentential clauses in the utterances. To illustrate these possibilities, let us consider the utterance and its representation below:

Bill: I am sick, and I believe that I am going to faint.

```
(SAY
  (BILL
   TIME-1
   JOHN
   (AND
        (IS BILL SICK TIME-1)
        (BELIEVE
                (BILL
                 TIME-1
                 (FAINT
                        (BILL
                         TIME-2]
```

The entire representation, [SAY ...] is a proposition in the first sense. Propositions in the second sense include (a) [SAY...], (b) [AND...], (c) [IS...], (d) [BELIEVE...] and (e) [FAINT...]. Propositions in the third sense include only (b)-(e), and not (a).

At the present stage of discussion, I can only throw out these alternatives for the sake of debate.

I.2.2.2 The team currently has a clearer notion of just what it is that they wish to explore under the heading of the term "repeated propositional reference" than they do of what they wish to define their notion of proposition as. So let us now turn to a consideration of some dialogue phenomena which the team (a) feels are instances of repeated propositional reference, and (b) are interested in investigating.

The following is a series of constructed examples of dialogue excerpts which the team would feel involve instances of repeated propositional reference phenomena of a type currently worthy of analysis.

REPEATED TEXT AND PROPOSITIONAL REFERENCE: CONCEPTS AND DETECTION
THE NOTIONS OF TEXT AND REPEATED PROPOSITIONAL REFERENCE

(1) A: Mary is sick.
B: That's unfortunate.

(2) A: I have no money. Five members of my family are deathly ill. I can't sleep at night because I have to keep fighting off the blood-thirsty rats in my vermin-infested apartment block.
B: Your story is not so different from the stories most other people in the neighborhood could tell.

(3) A: My daughter is expecting a child.
B: Yes, but it's a great secret; don't tell anyone else about it.

(4) A: John said that Mary is sick.
B: Paul told me that piece of bad news yesterday.

(5) A: John said that Mary is sick.
B: If what he said is true, we can't have our picnic.

(6) A: John believes that Mary is sick.
B: Yes, but Paul doubts it.

(7) A: John believes that Mary is sick.
B: What he believes is true, unfortunately.

When one first rapidly glances at the above examples, one feels that they are similar in that in each dialogue some pronoun or noun phrase is used to talk about something which has previously been talked about by the use of some sentence(s) or nominalized sentential phrase. One also feels that this similarity can only be specified in very vague terms, as was just done. And a closer look at the examples shows why: there is an extraordinary variety of things going on. In (1), the first sentence is used to describe a state of affairs, which is then referred to and commented on by the second utterance. In (2), the first turn involves a description of a state of affairs, and the second utterance comments not so much on that state of affairs as on the description which was made of it (even though one would clearly hesitate to say that a text reference was being made). In (3), the first turn involves the imparting of a piece of information which is a description of a state of affairs, but which is commented upon as a piece of information in the second turn (the information is true, but its a secret - i.e. it is a piece of information which has not been imparted to many people). In (4), indirectly quoted speech is reported, and then is said to be identical with some other reported speech; (5) again involves reported speech, but the statement which was said to be made is then treated as a proposition in a "transparent" context. In (6) and (7), an intentional object - a belief - is talked about, but in (6) it is talked about within an opaque context by both participants, whereas in (7) a belief is talked about once in an opaque context and again in a transparent context.

This heterogeneity is not such as to indicate that the PREF phenomena which the team currently wishes to study are completely ill-defined, however. First of all, the team

is well-united on the decision that they do not wish to study PREF phenomena which involve relationships between two synonymous sentences used in a dialogue; nor do they wish to study under the heading of PREF the relationship which holds between two utterances in a dialogue such that the participants who uttered them thereby "said the same thing" or made the same statement. Thus the following pairs of sentences, though they are instances of PREF, are not of current interest.

A: Mary is sick.
B: Mary is ill.

A: You are hot.
B: Yes, I am hot.

Secondly, there would seem to be some consensus on not including in the study of PREF an investigation of the relationship of co-referentiality between noun-phrases which denote what we would ordinarily regard as propositions or statements. NP-NP co-reference is thus outside of the scope of current PREF analysis, and such dialogues as those below are not to be studied under that heading:

A: John's assertion is simply not true!
B: I can't see why not. His claim seems well supported by all the available evidence.

A: De Morgan's law is a very important one.
B: Well, his theorem has certainly been useful.

Thirdly, there are certain referential phenomena which are, intuitively, quite distinct from what we are groping at above. One is reference to physical objects; another is reference to text per se. A third phenomena which is distinct is reference to actions, as in (8) and (9) below:

(8) A: John went fishing yesterday.
B: Mary did so too.

(9) A: Sky-diving without any training is exhilarating.
B: It's also foolish and suicidal. [14]

The distinction between propositional reference and action reference is often intuitively clear, but it is very hard to formulate. It cannot be pinned down in syntactic terms alone. One is tempted to say that neither the "it" nor the "so" in the above examples are propositional references because they cannot be analysed as standing for sentences, nominalized or no. The second turn in (8) could be rendered as "Mary did go fishing yesterday too" but not as "Mary did John went fishing yesterday too". However, the second turn of (9) could be rendered as "For people to go skydiving without any training is also foolish and suicidal". And the noun phrases or pronouns involved in PREF often

[14] An interesting example of reference to an action which can occur in dialogue is reference to a performative act previously performed by a participant, as in:
A: You're a bastard!
B: That's un-called for!

cannot simply be replaced by sentences: consider "your story" in example (2).

Given the above, we will say that repeated propositional reference of the type the team is currently interested in studying occurs in a dialogue when

(1) a sentential or multi-sentential utterance in the dialogue is subsequently mentioned by means of a pronoun or noun-phrase in such a manner that what is being mentioned is not the utterance considered purely as text (i.e. as an uninterpreted phonetic or graphic token, or as a member of some class of tokens).

(2) Some sentence uttered in the dialogue contains a nominalized sentence which refers to a state of affairs, a statement, a reported utterance or an intentional object, and some subsequent pronoun or noun-phrase is co-referential with that nominalized sentence.

Syntactically speaking, then, PREF involves only sentences and nominalized sentences on the one hand, and noun-phrases and pronouns on the other (see diagram overleaf).

The above delimitation of PREF is highly criticisable: it is a description by exclusion (cf (1)), and contains highly problematic terms (e.g. "state of affairs"). However, it is the least bad proposal I can come up with.

There is a terminological problem which remains to be dealt with. It is unsatisfactory to use the term "propositional reference", to describe a phenomena which covers some phenomena (those described in (1)) which do not involve reference in the usual sense at all. I shall continue using the term PREF, but only until a better term is found.

SECTION II:
DETECTION PROCEDURES
FOR
TEXT AND REPEATED PROPOSITIONAL REFERENCE

In this section, we will discuss constraints on TREF and PREF, as described above, which might be of some assistance in detecting their occurence by analysis of dialogue transcripts.

These constraints will be induced from an examination of dialogue samples in which TREF and PREF appear to occur. The examples will be of three types: some will be examples of real dialogue, some will be examples of dialogues drawn from literary works, such as plays or short stories, and some will be examples which have been made up (constructed) to illustrate a point or a difficulty. Each example provided will be marked by an R, an L or a C, according to its source.

The first observation which one makes when one considers actual dialogue is that TREF is much less frequent than PREF. Since TREF is more of an exceptional phenomena, we will examine it first, with the hope that occurences of it may be signaled more explicitly than are occurences of PREF.

II.1 THE DETECTION OF TEXT REFERENCE

As we have defined or described it above, TREF always involves the use of at least one expression to refer to text. A text-referring expression (TREXPR) [15] may be an ordinary noun-phrase ("his name", "what John wrote on the board"), or may be an expression which refers to text by either exhibiting or replicating it. So if we wish to detect instances of TREF, we must (a) find ways of determining whether an expression is being used as a TREXPR.

But once we have determined that a given expression is a TREXPR, we are not thereby assured that we have before us an instance of TREF as we have conceived of it. For a particular TREXPR, say E1, to be involved in TREF, one of the following two cases must obtain. (i) There is another token, say E, used in the dialogue, which is not a TREXPR, and which is either identical with or a member of the referent of E1. (ii) There is another TREXPR, say E2, used elsewhere in the dialogue, whose referent is either identical with, a subset of, or a member of the referent of E1. Thus, once we have detected the presence of a TREXPR in a dialogue (a), we must (b) compare its referent to other non-TREXPRs in the dialogue and to the referents of other TREXPRS in the dialogue, if there are such, and (c) decide on the basis of this comparison whether a TREF is occurring, and if so, what other expression is involved.

The above suggests an outline of a procedure for detecting TREF:

1. Find all occurences of TREXPRS.
2. For each TREXPR:

[15] For the remainder of section II, unless otherwise specified, we will use the term "expression" to mean expression-token.

2.1 Draw up a list of candidates for comparison with the TREXPR. This list will include
- other TREXPRs, if present
- "suitable" tokens, not TREXPRs, which are used in the dialogue.

2.2 Compare the TREXPR with the candidates.
- If the candidate is a TREXPR, determine whether its referent is identical with, a subset of, or a member of, the referent of the TREXPR under study.
- If the candidate is a non-TREXPR, detemine whether it is identical with or a member of the referent of the TREXPR under study.

If no comparisons suceed, conclude that no TREF is occuring.

If only one comparison succeeds, return the pair of successfully compared expressions as the TREF which is occuring.

If several comparisons succeed, continue.

2.3 Apply some evaluative criterion (or criteria) to the pairs of expressions which have been successfully compared. If one pair is clearly a "best" match, return it as the TREF; if several pairs are almost equally "good", return the TREXPR under study along with all other members of these pairs as the (multiple) TREF which is occuring.

With this vaguely defined procedure in mind, let us turn to a consideration of a corpus of examples of text reference (listed overleaf).

The corpus consists of dialogue examples which are either real or literary (fictional). The real examples are either examples of written dialogue or of oral dialogues which were subsequently transcribed; the literary examples were written. We will give less weight to confirmation of procedures by transcribed examples when those procedures rely upon orthographic cues (such as the presence of quotation marks or capitalization). The reason for this is that orthographic cues are furnished not by the original participants in the dialogues but by the transcriber, and are the result of a decision by the transcriber that some expression was being used as a TREXPR; to rely upon orthographic cues in such cases is a 'cop-out' from the point of view of someone who wishes to to specify computer programs which will detect TREF independently of human judgments.

We will consider examples in turn, proposing subprocedures or criteria for each which will then be applied to subsequent examples. We will at first only deal with noun phrases, which present - on the whole - fewer problems for analysis, and then go on to deal with pronouns.

Let us begin with example (1) below. (In all of our examples we will italicise the expressions which we feel intuitively are involved in a TREF, and number them for ease of subsequent discussion.

L: Yes, I linked to PARC-MAXC and transferred a file - but couldn't run it at PARC-MAXC - *IDFAS?* (1)

O: What's *IDFAS?* (2)

L: Sorry, mistyped - ideas?

(REAL - WRITTEN - SOURCE: OC32.PROTOCOL)

We intuitively perceive the second occurence of "IDFAS", (2), as a TREXPR. This suggests that we regard expressions which are not in our lexicon either as words of English or as names as TREXPRs; but this principle is wrong, because it would lead us to mark the first occurence of IDFAS, (1), as a TREXPR also, wheras we clearly percieve it to be a simple expression (albeit a mistyped version of an expression) which L uses. Similarly, simple capitalization is not a cue, for both occurences of "IDFAS" are capitalized. What seems to indicate to us that (2) is a TREXPR is not only that it is an unrecognized symbol, but that it is the subject of a question. So this leads to the formulation of a principle of TREXPR detection (TD):

(TD1.1)

If an expression is not in the lexicon, then
if it is the subject of a question,
it is a TREXPR.

This principle in the case of example (1) allows us to conclude that there is one and only one TREXPR. What are candidates for comparison with it? Intuitively, we perceive that there is only one non-TREXPR candidate: the first occurence of "IDFAS", (1), which is a token of the type of the TREXPR (2) (this is an instance of TREF by replication). This leads us to formulate the following principle of candidate selection (CS):

(CS1.1)

If there is a non-TREXPR expression in the dialogue
which is orthographically identical with the
TREXPR under study, then it is a candidate.

Since there is only one candidate, our procedure returns the two occurences of "IDFAS" as a TREF.

Let us now consider example (2) below:

L: Guess what ... that didn't work either. It took "*TTY*" (1) to be a filename.

O: Did you say *TTY* (2) or *TTY:* (3) ?

L: Just *TTY*. (4)

O: If you append the colon, then it will be recognised as a device designator instead of file name (I hope).

What are the TREXPRs here? Principle TD1.1 picks out expressions (2) and (3). But intuitively we know that the quotation-expression (1) is a TREXPR, because it is explicitly quoted. Wo we have a new TD principle:

(TD2.1)

If an expression is enclosed in quotation-marks,
then the quotation-expression (i.e. the
quotation-marks plus what they enclose)
is a TREXPR.

We still do not have sufficient TD principles, however, for intuitively we perceive the occurence of "TTY" in L's second turn (4) to be a TREXPR also, and neither TD1 nor TD2 would mark it as such. I feel that I recognise that token to be a TREXPR for two reasons: (a) it is an expression which is not in the lexicon, and (b) it is said to have been uttered by someone, viz. L (L's utterance is elliptical, but expandable into "I just said TTY.") In all of the constructed sentences below, one would detect a TREXPR:

John said/is saying/says/will say blurpagg.
Mary shouted/is shouting/shouts/will shout ARRGGHH.
I wrote/am writing/write/will write Xuytmon.

Note that linguistic-reception verbs have the same effect as these linguistic-production verbs ("hear" as well as "say"). So we modify TD1 as follows:

(TD1.2)

If an expression is not in the lexicon, then
if it is the subject of a question,
it is a TREXPR;
if it is the object of a verb of linguistic
production or reception,
then it is a TREXPR.

So now our principles allow us to recognize four noun-phrase TREXPRS: (1) in turn 1, (2) and (3) in turn 2, and (4) in turn 3. Applying CS1.1, we find that there are no non-TREXPR candidates. So we are left to take each of the four TREXPRs in turn and compare them with the three others.

We percieve (4) to be co-referential with both of the other occurences of "TTY", and we formulate the following principle of co-referentiality of TREXPRS (CR) to account for that fact:

(CR1.1)

If there are two TREXPRs, E1 and E2, and one of
them, say E1, is a quotation-expression, then
if E2 is orthographically identical with
the quotation-content (i.e. the string
between the quotation-marks) of E1,
then E1 and E2 are co-referential.

We perceive the first and second occurences of "TTY", (2) and (4), to be co-referential, which can be explained as follows:

(CR2.1)

If there are two TREXPRs, E1 and E2, and they are
orthographically identical,
then they are co-referential.

We also perceive each of the occurences of "TTY" to be co-referential with each other and with ""TTY"". This can be seen as a logical consequence of the fact that co-referentiality is an equivalence relation. One would be wrong in so concluding, however, for one must

remember that we are counting as co-referentiality in matters of TREXPRs (a) identity of referent (b) subset-superset relations of referents, and (c) member-set relations of referents. And of course, if x is a subset/member of a set z, and y is also a subset/member of z, it does not follow that x and y are either identical or bear subset-superset/member-set relations to each other. Nevertheless, I will temporarily adopt the following heuristic, with full knowledge that it is false, but with an eye to the fact that its falsity may not be revealed in most dialogues.

(CR3.1)

For any three TREXPRs, E1, E2, and E3, if E1 and E2 are both co-referential with E3, then E1 and E2 are co-referential with each other.

Lastly, we do not perceive (3) to be co-referential with any other noun-phrase TREXPR. This follows from the above CR principles.

In example (2) above, we relied upon the presence of verbs of linguistic production and reception to detect TREF. Certain nouns may also signal possible TREFs. Consider example (3) below:

O: Yep, and I will do my best to help. What is it you want? [name 1] Go ahead.

L: I would like to unarchive tapes 1120 and 1121 programs are called */name 2/* (1) ...~~~~~~~(interrupt here)~~~~~~~~

.... We would like to unarchive these.

O: In?

L: We are in directory [name 3] but the tapes were archived from the [name 4] directory. Go ahead.

Ok, but you will have to give me *those names* (2) again...

(REAL - WRITTEN - SOURCE: OC636.PROTOCOL)

We perceive expression (2) to be a TREXPR because of the meaning of the word "name": a name is a type of sound or inscription which we use to refer to individuals. So we adopt the following TD heuristic:

(TD2.1)

If an expression is a member of the set of conventional TREXPRS (C-TREXPRS), then it is a TREXPR.

(TD3.1)

The set of C-TREXPRS is <name(s)>.

We also perceive the expression (1) to be a TREXPR. The reason for this at first seems that it is the indirect object of the verb "call", which suggests the following general principle:

(TD4.1)

If any expression is the indirect object of the verb "call", then it is a TREXPR.

The principles TD3 and TD4 seem to be borne out by the following example (4):

A: Our president calls us "*Ambassadors of Friendship*". (1)

B: *Beautiful motto.* I winder if you know a conductor by the name of *George Whipple* (3) ?

A: George Whipple? No. I knew a *George Galloway* (2).

B: This is George Whipple.

A: I don't recall *the name* (4).

(LITERARY - WRITTEN - SOURCE: THE TRAVELLOR)

We can see that (3) and (4) are TREXPRs using TD4.1. (4), of course, requires semantic processing to establish that "by the name of X" introduces a textual object, X, which is a name.

We have been concentrating so far on noun-phrases, trying to decide whether they are TREXPRs, and what other noun-phrases or expressions TREXPR noun-phrases are associated with in occurences of TREF. Given the presence of "this" in B's second turn in example (4), we should begin to consider examples of TREF which involve pronouns. Here it is useful to consider certain general rules for finding the antecedents of pronouns which are useful in cases of repeated reference not only to text, but also to actions, objects and propositions. We shall only be concerned here with general rules for the determination of the antecedents of the pronouns "it", "that" and "this", given the fact that text is never referred to by the pronouns "he" or "she", because of gender considerations.

(GR1.1) IF A PRONOUN = "IT", "THAT" OR "THIS":

(1) If there is another previous pronoun of the same type in the same sentence, then that pronoun is a candidate of priority 1 for co-referentiality;

(2) If there is a pronoun in the nth preceding sntence (where n is less than some parameter I), then that pronoun is a candidate of priority n+1 for co-referentiality.

(3) If no pronominal candidates are found, then preceding noun-phrases are candidates.

(a) Noun phrases within same sentence have a higher priority than noun phrases in preceding sentences.

(b) Noun phrases in nth preceding sentence have lower

priority than noun-phrases in mth preceding sentence, where n and m are both less than some parameter I and n≤m.

(c) Within a sentence, noun phrases in subject position have a higher priority than noun phrases in object position; noun phrases in object position in turn have a higher priority than non phrases in prepositional phrases.

(d) Within a sentence, noun-phrases within a main clause have a higher priority than noun-phrases within a subordinate clause.

(e) Within a sentence, focused noun-phrases have a higher priority than noun-phrases which are not focused. (Focused noun-phrases are those moved to the front of sentences by such transformations as extraposition or tough-movement.)

(GR2.1)

Candidates for co-referentiality are to be selected by comparing the following features or predicates of the pronoun with those of the candidate: (a) sex and number (b) case or type constraints. If such considerations fail to select one candidate, rely on more specific plausibility considerations. [16]

Let us see how these general rules, in conjunction with the rules we have specified so far, allow us to detect TREF. Consider first the following example (5):

A: You've surely heard me speak of *Eugene Tesh* (1)!

B: I can't say that I have.

A: Well, *his name* (2) is always in the newspapers; he's a dramatic star. Everyone I know would recognize *it* (3).

B: I am not familiar with *the names of dramatic stars* (4). I have never seen *it* (5) before.

(CONSTRUCTED)

[16] These general rules were suggested by Jim Levin, on the basis of his own work and the work of others in the AI field. I do not claim that he would endorse the form that I have given them here.

According to our present rules, "Eugene Tesh" is marked as a TREXPR because it is not in the lexicon. When we get to expression (2), which is marked as a TREXPR by rules TD2 and TD3, (1) is the only candidate for partnership in a TREXPR. Now we understand that when reference to an individual is made by the introduction of his/her name, two pieces of information are imparted: (a) that there is an individual names X, and that there is a textual object, viz. X itself, which bears the relation name-of to that individual. And in this particular case, we understand that (2) is a reference to the name "Eugene Tesh", and not to the individual named Eugene Tesh. We might capture this by the following crude rule:

(CR4.1)

If there are two TREXPRs, one of which is or contains the word "name" or "names", and the other one of which has been interpreted as referring to an individual by name, then they are co-referential.

We are supposing that the phrase "interpreted as referring to an individual by name" has some meaning in terms of syntactic and semantic programs yet to be specified. When we come to the pronoun (3), the general rules comes into play. According to these rules, the first expressions which would be considered as candidates for antecedents of (3) would be noun-phrases, since there are no "it"s, "that"s or "this"s in the preceding sentences. Noun-phrases having higher priority than (2) would be rejected because of sex or number constraints (e.g. "a dramatic star"). And (2) would be selected. When we come to (4), CR4 would mark it as co-referential with (1), and CR3 would mark (1), (2), (3) and (4) as co-referential. Lastly, the general rules would mark (2) as the antecedent of (5), and CR3 would again chain (1)-(5) together as co-referential.

Another example in which our present rules would give us a satisfactory result would be the following example (6):

A: Fine! The other one is a *L.V.N.* (1) down at Permanente on Sunset Blvd.

B: *L.V.N.* (2) ... *That's* (3) lanky vertiginous nurse?

A: Right.

(REAL - ORAL - SOURCE: BLIND-DATES.PROTOCOL)

Our rules would mark (1) and (2) as TREXPRs, by TD1. We perceive (1) and (2) to be co-referential, by:

(CR5.1)

If two TREXPRs are not in the lexicon, then
if they are orthographically identical,
then they are co-referential.

When we come to the pronominal expression (3), the general rules would mark (2) as being the first noun-phrase acceptable candidate for being (3)'s antecedent. Note that this conclusion is very fragile. If B had not repeated "L.V.N.", the general rules would have selected "Sunset" as (3)'s antecedent.

A furthur example showing the additional development of the general rules is the following, example (7):

A: I wonder if you know a conductor by the name of *George Whipple* (1) ?

B: George Whipple? No. I knew a *George Calloway* (2).

A: This is George Whipple.

B: I don't recall *the name* (3).

(LITERARY - WRITTEN - SOURCE: THE TRAVELLOR)

Our rules mark (1) as providing a TREXPR; as above, we shall assume that the processor will, in the interpretation of A's first turn, (i) introduce an individual named George Whipple and (ii) introduce a textual object, "George Whipple", which bears the relation name-of to that individual. Similarly for (2). But now a difficulty becomes apparent. The pronoun "this" in "This is George Whipple" clearly refers to the individual referred to in the first turn, and not to his name. Our general rules would probably bind "this" to "George Galloway", and if they did not (on general grounds of the implausibility of the resultant conclusion that George Galloway is George Whipple), they would bind "this" to the name "George Whipple".

The above gives the flavor of the process of gradual development which might well lead to some acceptable rules for detecting TREF. Obviously, what has been said above only represents the beginning of such a process. I would like to end this section by mentioning some of the difficulties which the construction of TREF rules will undoubtedly encounter.

First of all, there are instances in which people make spurious use of orthographic cues of TREF. Consider example (8):

A: Hello. Got a couple of questions about "runoff" to on-line. Go ahead.

B: Okay. I've got a manual here, and although I don't know too much about it, we'll see what I can find. Hold a sec ... O.K. Shoot.

A: I have a rather old manual and I am trying to get runoff to print to my TTY on line...

(REAL - WRITTEN - SOURCE: OC370.PROTOCOL)

The use of quotation-marks around "runoff" in the first turn might well be regarded as a simple mistake (though there is an alternative interpretation, to which we will return below). If it is so regarded, the question arises: should one build rules that will test for and eliminate mistakes?

Secondly, there is the problem of TREXPRs which contain variables. An example would be "When making a new connection type ATT LINKER (PASSWORD), job number". Here only the "ATT" is a replication of its referent; "LINKER" stands in for a variable string, or a variable class of strings. Logicians have wrestled with the problem of variables within quotation contexts because they needed to be able to have variables ranging over text when constructing truth definitions; there are considerable problems here.

Thirdly, there is the problem of the use of quotation-marks in irony, the use of so-called snigger quotes. Jim Levin has suggested what is undoubtedly the right approach to such problems: regard quotation-marks in general as a signal that some peculiarity in processing the quoted words is required. The most frequent peculiarity thus signaled is that the words themselves should be retained, but other peculiarities should be allowed for. For example, quotation-marks may signal that a word is being used in a sense very different (perhaps opposite) from that in which it is commonly used, or that some presupposition of the use of that word is not obtaining. To return to example (8) above, for instance, it is possible that the user was quoting "runoff" to show that though the term usually is supposed to denote a program that runs off formatted copies, he is unwilling to use it in that way, because his experience leads him to believe that the program in question obstinately refuses to run off copies!

Despite these difficulties, I believe that one might well be able to devise a set of heuristics which would detect TREF correctly in a satisfactory number of cases. TREF is on the whole much less intractable than PREF, to which we now turn.

II.2 THE DETECTION OF REPEATED PROPOSITIONAL REFERENCE

Our general approach to the detection of repeated propositional reference will be largely similar to our approach to the detection of TREF. That is, we will first look for certain proposition-referring expressions, PREXPRs, and then, once PREXPR(s) have been found, look for other expressions which may be involved with those PREXPR(s) in repeated propositional reference. Lastly, once we have a PREXPR and a list of "candidates", we will select candidates by some criteria.

There will be important differences, however. PREF involves (a) a noun-phrase or pronoun on the one hand, and (b) a sentence or nominalised sentence, on the other. Now the first conclusion one reaches when one considers instances of PREF is that ***just about any sentence or nominalized sentence is capable of participating in a PREF.*** As a result, it is not functional to first pick a sentence or nominalized sentence and then look for a noun-phrase or pronominal candidate - there would simply be too much useless processing involved in such a procedure. Rather, one must first find a noun-phrase or pronoun which, because of its meaning, grammatical position or features (i.e. because of what is predicated of it) is suceptible of being involved in a PREF, and then look for sentence or nominalized sentence candidates.

Below, we will first analyse examples of PREF which involve noun-phrases, and then look at examples of PREF which involve pronouns.

First, let us consider example (1) below:

A: You know, I just... *The second question* (1) would be: *why wasn't this done before I went through all these bone scans, thyroid scans, and you know...* (2)

B: Well, that's a...

A: I'll die of radioactivity.

B: Yes, that's *the logical question* (3) and...

(REAL - ORAL - SOURCE: MEDICAL-CENTERED.PROTOCOL)

We see intuitively that "question" is a noun-phrase which refers to what we would call a proposition, so that both (1) and (3) are noun-phrases which might be involved in a PREF. Other noun phrases are similar to "question" in this respect, for example "statement", "request", "order", "demand", "query" and so on. This suggests the following principle for detecting PREXPRs (PD):

(PD1.1)

If an expression is a member of the set of conventional proposition-referring expressions (CPREXPRs), then it is a PREXPR.

(PD2.1)

> The set of CPREXPRs is <question(s), request(s), order(s), demand(s), assertion(s), claim(s), query(ies)....>

We also see that (2) is co-referential with both (1) and (3), because it is an expression of the interrogative form. What we are relying on here is a syntactic requirement which derives from the meaning of the noun "question". Certain other CPREXPRS impose syntactic requirements upon candidates, for example, "advice" in the following example (2):

> A: And now I shall give you *an extra piece of advice* (1). *Stop disgracing your daughter with your company on the streets - and, above all, at the theatre...* (2) or she will soon have every door to advancement shut to her!
>
> (LITERARY - WRITTEN - SOURCE: MOTHERLOVE)

"Advice" is a CPREXPR, and we know that (2) is a candidate for PREF because it is in the imperative mood. A third example of a dialogue in which syntactic cues deriving from the meaning of a CPREXPR are used in determining PREF is (3) below:

> A: General, I only want to keep one little private letter. Only one. *Let me have it.* (1)
>
> B: Is that *a reasonable demand* (2), madam?
>
> (LITERARY - WRITTEN - SOURCE: THE MAN OF DESTINY)

Here again, we see that (2) is co-referential with (1) because demands are (usually - I am being sloppy here of course) expressed by means of imperative sentences. One may therefore adopt the following heuristic:

(TD3.1)

> Given a CPREXPR, if that CPREXPR denotes a type of proposition which is usually expressed by a sentence of a certain grammatical mood (declarative, interrogative, imperative, horartory, etc.), then any nearby sentence or clause which is of that mood is to be considered as a PREXPR cadidate.

We have included the phrase "sentential clause" in TD3.1 because we want to account for cases like "Sam is curious to know whether or not the Socialists will take over the Portugese government. - That's a good question". In such cases one finds that embedded questions participate in PREF.

Many instances of PREF involving noun-phrases require a pretty complete understanding of the meaning of the noun-phrase in question in order to select candidates. I will give three examples below. First, example (4):

> A: Just a passing *comment* (1), Joe. We're having lunch right now, and I just made myself a hotdog sandwich with catsup. *Very tasty and almost unheard of in the old days.* (2)

(REAL - ORAL - SOURCE: CORRECTION-ACTIONS-CORPUS.PROTOCOL)

(I will suppose that the ellipsis in (2) has been filled out.) In deciding that (2) rather than the preceding sentence is a candidate for PREF involvement with the CPREXPR (1), we rely upon our understanding of what a comment is, i.e. an observation or remark expressing an opinion or attitude. Similarly, in example (5) below:

A: ... And, Mary, I can tell you *a secret.* (1) It's still *a great secret* (2), mind! *They're expecting a grandchild.* (3) Isn't that *good news* (4)?

(LITERARY - WRITTEN - SOURCE: THE LONG CHRISTMAS DINNER)

our selection for candidates for (1) and (2) and for (4) depends (a) on our knowledge of what a secret and what news is, and (b) on what kind of information would probably count as a secret or as news to the participants given the situation and the participants' knowledge. Lastly, consider example (6):

A: So, anyway when we got there *the funniest thing* (1) happened. *They sat down and they passed out these little booklets* (2), because we went to their suite.

B: Uh-huh.

A: *And, they started preaching about their religion the whole three hours* (2) and we were just crawling the walls to get out.

(REAL - ORAL - SOURCE: BLIND-DATES.PROTOCOL)

This is a complex example. First, we recognize (1) as being a PREXPR, not because it is itself a CPREXPR, but because it is the subject of a verb which takes as subjects nouns which refer to events, e.g. the verb "happen". This leads to a new principle:

(PD3.1)

If an expression is the subject of an event-verb,
then it is a PREXPR.

(PD4.1)

The set of event-verbs is <happen, occur, ...>

We also know that candidates for involvement in PREF with event-nouns must be declarative sentences which describe events or states of affairs, e.g. they cannot express general laws. I hesitate to make a rule of candidate selection out of this intuition, however, because I can think of no operationalizable way of detecting when a declarative sentence describes an event or state of affairs. (This requires furthur work, to say the least!) But let us return to our previous comment about the need to understand the meaning of PREXPR-nouns in order to select candidates. We intuit that it is not a strange or funny thing that the speaker and the group of people the speaker was with "started

crawling the walls to get out", given the circumstances. It is this complex understanding which allows us to decide that the last part of the third turn is not part of the PREF. We also intuit that the last clause of the first turn is not part of the PREF, and we do so not only because we understand that, in the circumstances described, it was not a strange thing for the speaker to go to their suite, but also because of our understanding of the semantic function of the clause "because we went to their suite". There are two different functions of clauses prefaced by the word "because", illustrated by the two following sentences:

(a) John is not coming to the meeting tonight,
because he is sick.

(b) John is not coming to the meeting tonight,
because he just phoned me from Australia.

"Because" can either be used to talk about causes, as in (a), or to introduce considerations which either logically or plausibly justify making a certain statement, as in (b). In the example above, "because" is being used to explain or justify the making of a certain description, and is thus not part of that description itself.

Let us now turn to an examination of some cases of PREF which involve pronouns.

Some general heuristics may be laid down at the outset. The first concerns the distinction between pronouns which refer to actions and pronouns which refer to propositions. Consider the following example (7):

A: And, for your information, Jack, I'm just going ***to tear into some beef and gravy and other assorted goodies.*** (1)

B: I presume that you're doing ***this*** (2) with the full permission and - of the commander.

(REAL - ORAL - SOURCE: APOLLO-13/PAGE379.PROTOCOL)

Here we intuit that the pronoun (2) is involved in what might be called a repeated reference to an action. The principle clue is that (2) is the object of the pro-verb "do". This clue, in so far as I have been able to ascertain, is a frequent and reliable one. It is not the only clue, however, as the example (8) shows:

A: We have decided to use a canister and, you know that the liquid-cooled garment has a bag around it that we think we can use too, or that we know we can use. We've tried ***it*** (1).

(REAL - ORAL - SOURCE: APOLLO-13/PAGE379.PROTOCOL)

Here we intuit that (1) is involved in a repeated action reference, although just what action is involved is rather unclear. The clue here is that the pronoun (1) is the object of the verb "try". So we can formulate the following rather solid heuristic:

(PD5.1)
If a pronominal expression is the object of either of

the verbs "do" or "try", then it is not a PREXPR.

Secondly, we know that there are certain predicates which mark pronouns as PREXPRs. Some of these can be grouped into two classes: the class of what can be loosely called "logical" predicates, such as "x is true/false" , "x is possible/probable", "x is inconsistent" or "x implies y", and what can be called intentional predicates, such as "x knows/believes y".

(PD6.1)

If an expression is such that some logical predicate is attributed to its referent, then it is a PREXPR. Logical predicates include the adjectives "true", "false", "probable", "possible", and the verbs "imply", "entail".

(PD7.1)

If an expression is the object of an intentional verb, then it is a PREXPR. Intentional verbs are a class of verbs which includes "know", "believe", "remember", "wants".

As soon as one thinks about PD7.1, however, one realizes that it is insufficient. Most intentional verbs [17] can take expressions as objects which denote not propositions, but objects. Thus we have not only "John remembers that Mary is sick and Paul remembers it too", but also "John remembers Bill's boat and Paul remembers it too". An example of such a use of "know" is the following:

A: ... And they started driving and I don't know if you know San Gabriel Valley where Crystal Lake is?

B: I don't know the area too well, my dear.

(REAL - ORAL - SOURCE: BLIND-DATES.PROTOCOL)

However, if one restricts PD7 to pronominal expressions, then it can be defended on the basis of a frequency argument. If one examines the occurences of the verbs "know" and "believe" followed by "it" or "that" in the dialogues which are presently on line, one finds that the great majority are instances of PREF phenomena. A case by case study of intentional verbs is required here. But for the moment, let us amend and restrict PD7:

(PD7.2)

If a pronominal expression is the object of one of the two intentional verbs "know" and "believe", then it is a PREXPR.

One last clue that we can propose for the detection of pronominal PREXPRs is that the pronoun "so", when it is the object of a verb, is always a PREXPR (as far as I have been able to determine). Examples are: "Is John sick? - I think so." and "I hope that Mary passed her exam. - I hope so too."

(PD8.1)

All occurences of the pronoun "so" as the objects of

[17] On intentional verbs and their logical peculiarities, see the appendix on intentionality.

verbs other than the pro-verb "did" are PREXPRs.
This clue is infrequent but very reliable.

Let us now turn to the consideration of some examples of PREF which involve pronouns. Example (9):

A: Well, ***you might have saved your life*** (1), my dear Caty.

B: I know ***that*** (2).

A: Yeh, ***you might all have been wiped out in a drunken car accident if you hadn't done that.*** (3)

B: I know ***it*** (4).

(REAL - ORAL - SOURCE: BLIND-DATES.PROTOCOL)

Both (2) and (4) are marked as PREXPRs by rule PD7.2. The problem of candidate detection and selection now arises. We perceive (1) to be co-referential with (2) and (3) to be co-referential with (4). This suggests the following pair of blatantly rudimentary rules of candidate detection (PCD) and candidate selection (PCS):

(PCD1.1)
The candidates for co-referentiality with a pronominal PREXPR is the set of all sentences and nominalised sentences at a distance of m sentences from the PREXPR in question (before or after), where m is some search parameter.

(PCS1.1)
Select the first preceding sentence or nominalized sentence as being co-referential with a pronominal PREXPR.

These two rules seem to work in our next example (10):

A: Very briefly, I've had a lot of pain for six weeks and diagnosed more or less as a dislocated disc. Now, what's your feelings? In the first place, they said cancer - maybe. So, I had a lot of tests done, but now, since I changed doctors, he says there is, well, ***there's this blood test called C.E.A., which will tell if there is cancer anywhere in your system*** (1). And I could hardly believe ***it*** (2). Now, is there such a test?

(REAL - ORAL - SOURCE: MEDICAL-CENTERED.PROTOCOL)

And again in the following examples, drawn from the same source and which we will bring together as example (11):

A: My wife knows I wear my cap the way I like to. And I know what's good for my wife, as well as for everybody else. ***I keep my business to myself, without any need of those who wear feathers in their caps.*** (1) And everybody in these parts knows ***it*** (2), thank the Lord!

A: ***Her husband arrived only this morning.*** (1)

B: Oh, oh, you know ***that*** (2) too? Bravo!

A: ***Your're out of your mind!*** (1)

B: Yes, ***it's*** (2) true! I'm out of my mind!

A: I'm going home because my husband's on my mind. ***I didn't see him in church.*** (1)

B: Don't think of ***that*** (2). He'll be along to the square.

(LITERARY - WRITTEN - SOURCE: CAVALLERIA RUSTICANA)

These rules will also operate satisfactorily on the following example (11) if supplemented by the general principles for pronoun resolution GR1 and GR2 set forth above in section II.1:

A: I heard them say ***that she had been a loose woman*** (1)! I don't want to believe ***it*** (2) - I still don't believe ***it*** (3) - but I can't help feeling that ***it*** (4) is true. Everything points to ***it*** (5) - and I feel ashamed, mortified! Ashamed to show myself in her company. Everybody seems to be staring at us - I seem to feel the men ogling us! ***It's*** (6) frightful! But can ***it*** (7) really be true? Do you think ***it*** (8) can be true? Tell me!

(LITERARY - WRITTEN - SOURCE: MOTHERLOVE)

If the search parameter m of the general rules is sufficiently large, the expressions (2)-(8) would be determined to be co-referential with (1). Some disatisfaction might be felt with this result in regard to (6), however, since what is said to be frightful is perhaps the speaker's shame and embarassment rather than (or perhaps as well as) the purported fact that she (the mother) had been a loose woman. Intuitions are not very clear on this point, and the question can be answered either way with little impact on the dialogue analysis in

this particular case. The problem, though not very serious, signals a fact that we will return to below, viz. the fact that pronominal PREXPRs which have attitudinal adjectives predicated of them are more difficult to select candidates for than pronouns which have, say, logical predicates.

For the moment, let us comment furthur on PCS1.1 It is unclear just what we mean by "sentence" in that rule: do we mean literally a string of words ending with a period, or a sentential constituent, of which there may be several in a sentence in the literal sense? The latter interpretation seems required by examples such as the following (13):

A: Sorry to bother you, but *someone seems to have changed one of our passwords* (1) and no one here knows anything about *it* (2).

(REAL - WRITTEN - SOURCE: OC133.PROTOCOL)

and the following example (14):

A: Fred, *in about 4 minutes, we're going to hand you over to a different communications site, and it's going to take us about a minute or so to re-establish uplink* (1), so you can be prepared for *that* (2).

(REAL - ORAL - SOURCE: APOLLO-13/PAGE 379.PROTOCOL)

Secondly, we must ask how PCS1.1 fares when the preceding sentence is in the interrogative mood. There are some cases involving the PREXPR "so" in which that pronoun is involved in a PREF with the declarative transform of the preceding interrogative sentence: "Is Mary sick? - I believe/think so." Such cases only occur when the preceding question is not of the WH-type: consider the absurdity of "Who is the President? - I believe so." Apart from the special "so" cases, there seem to be two other kinds of cases, illustrated by (a) and (b):

(a) Is Mary sick?
I don't know that.

How many feet are there in a meter?
Oh, I learned that in school.

(b) Is Mary sick?
What makes you ask that?

How many feet are there in a meter?
I often wonder about that myself.

In cases of type (a), the pronoun stands in for an answer to the preceding question, whereas in cases of type (b), the pronoun stands in for the preceding question itself. It is difficult to find a principle which would distinguish between the two cases. One possible solution would focus upon the role which the verbs in the verb-phrases containing the pronouns usually play with regard to presuppositions. This is a problem I hope to do more work on; at present I can only pose it.

Thirdly, it is clear that PCS1 is biased in favor of backward pronominalization. It will give false results in cases such as the following example (16):

A: Just what are you trying to tell me, young lady?

B: What I want to tell you is *this* (1) : ***Your daughter has an opportunity to come out among people - and, perhaps, either to advance her career and gain recognition, or to become engaged to and marry a young man of a good, respectable family...*** (2)

(LITERARY - WRITTEN - SOURCE: MOTHERLOVE)

PCS1 should therefore be modified so as to allow for forward propositional pronominalization, at least by treating "this", as opposed to "it" or "that", as a cue for such a forward direction.

Lastly, there are clearly cases in which the rule according to which one should select the preceding sentential clause or nominalized sentential clause would fail, such as example (17) below:

A: *The other one is an L.V.N. down at Permanente on Sunset Blvd.* (1)

B: L.V.N. ... That's lanky vertiginous nurse?

A: Right.

B: I think *that* (2) 's terrific. Listen, what are you getting Sam for your 29th anniversary?

(REAL - ORAL - SOURCE: BLIND-DATES.PROTOCOL)

Such a case might be handled by modifying PCS1 so that those candidates which were involved in "correction-actions" (in a large sense) would not be selected. Such a rule would be difficultly operationalizable, however, and given the frequency with which the present focus on the preceding sentential clause or nominalized sentential clause proves itself to be useful, it should probably be retained.

Having sketched above the very beginning of procedures for detecting PREF in relatively tractable cases, I would like to list some of the difficulties found in more unmanageable cases.

The first major problem which I see turns on the fact that we can say many things about events, states of affairs, reported speech and intentional objects. Some comments, such as comments about whether or not we believe that they obtain, or about our judgments of the truth or falsity of propositions which convey information about them, clearly apply only to propositions. But many other comments could equally well be made about physical objects. And when such comments are made, they do not provide us with

any clear way of determining that we are faced with a PREXPR. This is especially true of attitudinal comments. Consider example (18):

A: One is out in the garage. ***She's a teacher, and she's got all her things out there and she's got her little radio there.*** (1)

B: Hey, *that* (2) 's neat. What's her first name?

(REAL - ORAL - SOURCE: BLIND-DATES.PROTOCOL)

The adjective "neat" can be predicated of objects and actions as well as of states of affairs. It requires much sophisticated processing to determine that (2) is probably co-referential with all of (1), rather than with, say, "her little radio". In particular, on relies on one's knowledge of the situation in which the communication is taking place, a situation in which the participants are not in the same location. Let us consider another example (19):

A: And, Aquarius, for your information, ***we now have 136-mile perigee.*** (1) Confirmed by Doppler.

B: Okay, ***136-mile perigee now.*** (2) *That* (3) 's very nice.

(REAL - ORAL - SOURCE: CORRECTION-ACTIONS-CORPUS.PROTOCOL)

We intuit that (3) is a PREXPR, and that it is co-referential with (2) and therefore with (1). However, it is rather difficult to know just how we do this. It would be rash to propose a rule which made such predicates as "is nice" signals of PREF, because such predicates are probably more frequently attributed to objects than they are to states of affairs. And so one is left with routines which would mark (3) as an object reference, co-referential with the noun-phrase "136-mile perigee".

The second major difficulty one should mention is that pronominal PREXPRs often have what one might call indefinite scope. That is, they are co-referential with a large but indeterminate number of preceding or subsequent propositions. As an example of this consider the expressions (1)-(6) in the lengthy example (19) below; all are indefinite in scope in differing degrees, and several involve forward propositional pronominalization (e.g. (1)):

C
...your worst experience on a blind date - especially if your first name is ...

T
Sharon

C
Sharon, how old are you?

T
I'm 21.

C
O.K., are you blond or brunette?

T
I'm a brunette.

C
O.K. Tell me about your worst experience on a blind date, my dear.

T
Well, my worst experience happened when I was uh, well I was in college, I was going to a girl's school.

C
uh, huh ...

T
...and they used to have all these singing groups, you know, come in and entertain us.

C
Do you mean singing groups - professionals, or do you mean from fraternities?

T
No, these were professionals singing.

C
Oh, yeh. Like the Four Freshmen and people like that?

T
Yeh, well, I ...

C
Like the Four Preps.

T
I'm not going to name the group.

C
Yeh.

T
Because *IT (1)* was really strange, because there were

about 3 of us and one girl knew the group. So, you know she asked would we like to go on this date with this singing group? We thought, Oh Wow, Yes!

C
You mean, two girls go out with four guys?

T
No, there were four guys but there would be 4 of us.

C
Oh, I see.

T
So, I got 2 other friends and you know, we thought *THIS (2)* was really going to be fantastic. We were going to get drunk and have a great big dinner and really have a ball.

C
Uh!

T
And so, uh, they had a limousine to come pick us up and it took us to the motel and we started, you know, giving each other the eye and getting kind of nervous.

C
They brought you right over to the motel. They figured you were going to sign up as "groupies".

T
Exactly, I think. That's what we thought, at least.

C
And, no dinner?

T
And no dinner, yes!

C
Wow!

T
So, anyway when we got there the funniest thing happened. They sat down and they passed out these little booklets, because we went to their suite.

C
Uh huh.

T

And, they started preaching about their religion the whole three hours and we were just crawling the walls to get out.

C
It was a religious frenzy?

T
Yeh!

C
Were they sitting around in their Saffron robes?

T
Exactly!

C
And their little Chinese Temple gongs?

T
Right!

C
What a weird, what a weird, wow!

T
And, we said, well can we order a drink? And, they said, Oh, *THAT's (2)* not the way of our religion - we can't drink.

C
Oh....

T
IT (3) was really the most boring time as compared to what we were expecting. I think it would have been more fun to go thrashing the hotel room.

C
Why of course! How long did you girls hand around with these religious freaks?

T
Well, we hung around ... I guess we left around 11 and I guess we got back around 3 in the morning.

C
On wow, you stayed too long, honey.

T
Yeh, but we were trapped actually.

C
Yeh, well did *THAT (4)* turn you against Saffron robes permanently?

T
Oh well, no, not really but I just wasn't expecting *it* (5) at that time.

C
Yeh. *THAT (6)* a dandy. I'm delighted you called, Sharon. You really surprised me. I thought I was going to have to bleep you out.

T
No. I called you twice before, Bill.

In conclusion, we may say that PREF detection is considerably more difficult than is TREF detection. It is hoped that the few preliminary approachs presented above are useful if only because suggestive.

BIBLIOGRAPHY

AUNE, Bruce, "Statements and Propositions", Nous, Vol. 1, August 1967

BAR-HILLEL, Y., "Primary Truth-Bearers", Dialectica, Vol. 27, Nos. 3-4, 1973

BARENSE, J.G., "Identity in Indirect Discourse", Journal of Philosophy, Vol. 66, June 19 1969

BERG, Jan, "What is a Proposition?", Logique et Analyse, Vol. 10, December 1967

CARNAP, Rudolf, "Meaning and Necessity", The University of Chicago Press, Chicago, 1956

CARTWRIGHT, Richard L., "Propositions", in 'Analytical Philosophy', R.J. Buttler, ed., Barnes Noble, New York, 1962

CARTWRIGHT, R., "Propositions Again", Nous, Vol. 2, August 1968

CHURCH, Alonzo, "On Carnap's Analysis of Assertion and Belief", Analysis, Vol. 10, 1950

CHURCH, Alonzo, "The Need for Abstract Entities in Semantic Analysis", Proceedings of the American Academy of Arts and Sciences, Vol. LXXX, 1951-1954

COHEN, L. Jonathan, "Geach's Problem about Intentional Identity", Journal of Philosophy, Vol. 65, May 30, 1968

CUSHING, Steven, "The Semantics of Sentence Pronominalization", Foundations of Language, Vol. 9, No. 2, 1972

DENNETT, D.C., "Geach on Intentional Identity", Journal of Philosophy, Vol. 65, May 30, 1968

FABER, Richard N., "Statements and What is Stated", Philosophical Studies, Vol. 23, Nos. 1-2, February 1972

FITCH, Frederic B., "The Reality of Propositions", The Review of Metaphysics, Vol. 9, 1955

FREED, Bruce, "Saying of and Saying That", Journal of Philosophy, Vol. 67, No. 22, November 19, 1970

GALE, Richard M., "Propositions, Judgments, Sentences and Statements", in 'The Encyclopedia of Philosophy', Paul Edwards, ed., Macmillan The Free Press, New York, 1967

BIBLIOGRAPHY

GARNER, Richard T., "Lemmon on Sentences, Statements and Propositions", Analysis, Vol. 30, No. 3, January 1970

GEACH, P.T., "Intentional Identity", Journal of Philosophy, Vol. 64, October 26, 1967

GEACH, Peter, "Assertion", in 'Readings in the Philosophy of Language', Jay ROSENBERG and Charles TRAVIS, eds., Prentice-Hall, Englewood Cliffs, New Jersey, 1971

GENOVA, A.C., "'Jonesese' and Substitutivity", Analysis, Vol. 31, No. 3, January 1971

GINSBERG, Mitchel, "Towards Statement Identity", Critica, Vol. 4, May-September 1970

GODDARD, L., and ROUTLEY, R., "Use, Mention and Quotation", The Australasian Journal of Philosophy, Vol. 44, No. 1, May 1966

GROVER, Dorothy L., CAMP, J.L., and BELNAP, Nuel D., "A Prosentential Theory of Truth", Philosophical Studies, Vol. 27, 1975

HINTIKKA, J., "Semantics for Propositional Attitudes", in Herbert FEIGL, Wilfrid SELLARS and Keith LEHRER, 'New Readings in Philosophical Analysis', Appleton-Century-Crofts, New York, 1972

HIZ, Henry, "Referentials", Semiotica, Vol. 1, No. 2, 1969

KATZ, Jerrold J., "Semantic Theory", Harper and Row, New York, 1972

LEMMON, E.J., "Sentences, Statements and Propositions", in "British Analytical Philosophy", B.A.O. Williams Alan Montefiore, eds., Routledge Kegan Paul, London, 1966

LINCICOME, D.V.C., "Systematically Ignored Differences and the Identity of Propositions", Foundations of Language, Vol 12, No. 1, September 1974

LITHOWN, R.J. MARRAS, A., "Intentionality Without Extensionality", Philosophical Studies, Vol. 25, August 1974

LYCAN, William G., "Eternal Sentences Again", Philosophical Studies, Vol. 26, Nos. 5-6, December 1974

MARRAS, A., "Intentionality and Cognitive Sentences", Philosophy and Phenomenological Research, Vol. 29, No. 2, December 1968

MORICK, Harold, "Intentionality, Intensionality and The Psychological", Analysis, Vol. 32, No. 2, December 1971

BIBLIOGRAPHY

NOLAN, Rita, "Truth and Sentences", Mind, Vol. 78, No. 312, October 1969

O'CONNOR, D.J., "Tests for Intentionality", American Philosophical Quarterly, Vol. 4, April 1967

PAGE, Edgar, "Reference and Propositional Identity", Philosophical Review, Vol. 79, January 1970 PAP, Arthur, "Propositions, Sentences and the Semantic Definition of Truth", Theoria, Vol. 20, 1954

PARTEE, Barbara Hall, "The Syntax and Semantics of Quotation", in 'A Festschrift for Morris Halle', Stephen R. ANDERSON and Paul KIPARSKY, eds., Holt, Rinehart and Winston, New York, 1973

PITCHER, George, "Propositions and the Correspondence Theory of Truth", in 'Readings in the Philosophy of Language', Jay ROSENBERG and Charles TRAVIS, eds., Prentice-Hall, Englewood Cliffs, New Jersey, 1971

POSPESEL, Howard, "The Non-existence of Propositions", Monist, Vol. 53, No. 2, April 1969

REINHARDT, L.R., "Propositions and Speech Acts", Mind, Vol. 76, April 1967

RICHARDS, Barry, "A Point of Reference", Synthese, Vol. 28, 1974

RYLE, Gilbert, "Are There Propositions?", Proceedings of the Aristotelian Society, Vol. 30, 1929-1930

SAYWARD, Charles, "Propositions and Eternal Sentences", Mind, Vol. 77, October 1968

SCHEFFLER, Israel, "An Inscriptional Approach to Indirect Quotation", Analysis, Vol. 14, 1954

SCHEFFLER, Israel, "Inscriptionalism and Indirect Quotation", Analysis, Vol. 19, 1958

SCHEFFLER, Israel, "Anatomy of Inquiry", Alfred A. Knopf, New York, 1963

SNYDER, Aaron, "On What Has Truth Value", Philosophical

SOSA, Ernest, "Propositional Attitudes de Dicto and de Re", Journal of Philosophy, Vol. 67, No. 21, November 5, 1970

STALNAKER, Robert C., "Wallace on Propositional Attitudes", Journal of Philosophy, Vol. 66, No. 22, November 20, 1969

THOMSON, James F., "Truth-Bearers and the Trouble about Propositions", Journal of Philosophy, Vol. 66, No. 21, November 6, 1969

TOMBERLIN, James, "Statements and Their Idenity Conditions", Logique et Analyse, Vol. 11, 1968

BIBLIOGRAPHY

WALLACE, John, "Propositional Attitudes and Identity", Journal of Philosophy, Vol. 66, No. 6, March 27, 1969

ZEMACH, Eddy M., "Reference and Belief", Analysis, Vol. 30, No. 1, October 1969

ZIFF, Paul, "What is Said", in D.DAVIDSON and G.HARMAN, eds., "Semantics of Natural Language", D.Reidel, Dordrecht-Holland, 1972

ZWICKY, Arnold M., "On Reported Speech", in C. H. FILLMORE and D. T. LANGENDOEN, "Studies in Linguistice Semantics", Holt, Rinehart and Winston, Inc., New York, 1971

DISTRIBUTION LIST

AC-S, Education Programs
Education Center, MCDEC
Quantico, VA 22134

AFHRL-AS (Dr. G.A. Eckstrand)
Wright-Patterson AFB
Ohio 45433

ARI Field Unit - Leavenworth
P.O.Box 3122
Ft. Leavenworth, KS 66027

Advanced Research Projects Agency
Administrative Services
1400 Wilson Blvd.
Arlington, VA 22209
Attn: Ardella Holloway

Air University Library
AUL-LSE 76-443
Maxwell AFB, AL 36112

Prof. Earl A. Alluisi
Code 287
Dept. of Psychology
Old Dominion University
Norfolk, VA 23508

Dr. Daniel Alpert
Computer-Based Education
Research Laboratory
University of Illinois
Urbana, IL 61801

Dr. John R. Anderson
Dept. of Psychology
Yale University
New Haven, CT 06520

Armed Forces Staff College
Norfolk, VA 23511
Attn: Library

Ms. Carole A. Bagley
Applications Analyst
Minnesota Educational
Computing Consortium
1925 Gather Ave.
Lauderdale, MN 55113

Dr. James Baker
U.S. Army Research Institute
1300 Wilson Blvd.
Arlington, VA 22209

Dr. M.A. Bertin, Scientific Director
Office of Naval Research
Scientific Liaison Group-Tokyo
American Embassy
APO San Francisco 96503

Dr. Jack R. Borsting
Provost & Academic Dean
U.S. Naval Postgraduate School
Monterey, CA 93940

Dr. John Brackett
SofTech
460 Totten Pond Rd.
Waltham, MA 02154

Dr. Robert K. Branson
1A Tully Bldg.
Florida State University
Tallahassee, FL 32306

Dr. John Seeley Brown
Bolt, Beranek, and Newman, Inc.
50 Moulton St.
Cambridge, MA 02138

Dr. Victor Bunderson
Institute for Computer Uses
in Education
355 EDLC Brigham Young University
Provo, UT 84601

Dr. Ronald P. Carver
School of Education
University of Missouri-Kansas City
5100 Rockhill Rd.
Kansas City, MO 64110

Jacklyn Caselli
ERIC Clearinghouse on
Information Resources
Stanford University
School of Education - SCRUT
Stanford, CA 94305

Century Research Corporation
4113 Lee Highway
Arlington, VA 22207

Chairman, Leadership & Law Dept.
Div. of Professional Development
U.S. Naval Academy
Annapolis, MD 21402

Chief of Naval Education and
Training Support (01A)
Pensacola, FL 32509

Dr. Kenneth E. Clark
College of Arts & Sciences
University of Rochester
River Campus Station
Rochester, NY 14627

Dr. Allan M. Collins
Bolt, Beranek and Newman, Inc.
50 Moulton St.
Cambridge, MA 02138

Dr. John J. Collins
Essex Corporation
6305 Caminito Estrellado
San Diego, CA 92120

Commandant
U.S. Army Institute of Administration
Attn: EA
Fort Benjamin Harrison, IN 46216

Commanding Officer
Naval Research Laboratory
Code 2627
Washington, DC 20390

Commanding Officer
Naval Health Research Center
San Diego, CA 92152
Attn: Library

Capt. H.J. Connery, USN
Navy Medical R&D Command
NNMC, Bethesda, MD 20014

Dr. T.E. Cotterman
AFHRL-ASR
Wright Patterson AFB
Ohio 45433

Mr. Joseph J. Cowan, Chief
Psychological Research Branch
(G-P-1-62)
U.S. Coast Guard Headquarters
Washington, DC 20590

DCDR, USAADMINCEN
Bldg. #1, A310
Attn: AT21-OED Library
Ft. Benjamin Harrison, IN 46216

Dr. Ruth Day
Dept. of Psychology
Yale University
2 Hillhouse Ave.
New Haven, CT 06520

Defense Documentation Center
Cameron Station, Bldg. 5
Alexandria, VA 22314
Attn: TC

Director, Offi of Manpower
Utilization
HQ, Marine Co. (Code MPU)
BCB, Building 2009
Quantico, VA 22134

Director, Management Information
Systems Office
OSD, M&RA
Room 3B917, The Pentagon
Washington, DC 20301

Dr. Donald Dansereau
Dept. of Psychology
Texas Christian University
Fort Worth, TX 76129

Dr. Ralph Dusek
U.S. Army Research Institute
1300 Wilson Boulevard
Arlington, VA 22209

ERIC Facility-Acquisitions
4833 Rugby Ave.
Bethesda, MD 20014

Dr. John Eschenbrenner
McDonnell Douglas Astronautics
Company-East
PO Box 30204
St. Louis, MO 80230

Major I. N. Evonic
Canadian Forces Personnel
Applied Research Unit
1107 Avenue Road
Toronto, Ontario, CANADA

Dr. Beatrice Farr
U.S. Army Research Institute
1300 Wilson Boulevard
Arlington, VA 22209

Dr. Marshall J. Farr, Director (4)
Personnel & Training Research Programs
Office of Navy Research (Code 458)
Arlington, VA 22217

Dr. James M. Ferstl
Employee Development:
Training Technologist
Bureau of Training
U.S. Civil Service Commission
Washington, DC 20415

Dr. Victor Fields
Dept. of Psychology
Montgomery College
Rockville, MD 20850

Dr. John Ford
Navy Personnel R&D Center
San Diego, CA 92152

Dr. Larry Francis
University of Illinois
Computer Based Educatonal
Research Laboratory
Champaign, IL 61801

Dr. John R. Frederiksen
Bolt, Beranek & Newman, Inc.
50 Moulton St.
Cambridge, MA 02138

Dr. Alfred R. Fregly
AFOSR-NL, Building 410
Bolling AFB, DC 20332

Dr. Frederick C. Frick
MIT Lincoln Laboratory
Room D 268
PO Box 73
Lexington, MA 02173

Dr. Vernon S. Gerlach
College of Education
146 Payne Bldg. B
Arizona State University
Tempe, AZ 85281

Dr. Robert Glaser
Co-Director
University of Pittsburgh
3939 O'Hara St.
Pittsburgh, PA 15213

Capt. D.M. Gragg, MC, USN
Head, Section on Medical Education
Uniformed Services Univ. of
the Health Sciences
6917 Arlington Road
Bethesda, MD 20014

Dr. Duncan Hanser
School of Education
Memphis State University
Memphis, TN 38118

Dr. Frank J. Harris
U.S. Army Research Institute
1300 Wilson Boulevard
Arlington, VA 22209

Col. Frank Hart, Director
Training Management Institute
U.S. Army, Bldg. 1725
Fort Eustis, VA 23604

Col. G.B. Howard
U.S. Army
Training Support Activity
Fort Eustis, VA 23604

HumRRO-Columbus Office
Suite 23
2601 Cross Country Dr.
Columbus, GA 31906

HumRRO-Ft. Knox Office
P.O. Box 293
Fort Knox, KY 40121

HumRRO-Western Division
27857 Berwick Drive
Carmel, CA 93921
ATTN: Library

Human Resources Research Organizatio
400 Plaza Bldg.
Pace Blvd. at Fairfield Drive
Pensacola, FL 32505

Dr. Earl Hunt
Dept. of Psychology
University of Washington
Seattle, WA 98105

Instructional Technology Branch
AFHRL
Lowry AFB, CO 80230

Dr. Edgar Johnson
U.S. Army Research Institute
1300 Wilson Boulevard
Arlington, VA 22209

Dr. Wilson A. Judd
McDonnell-Douglas Astronautics Co.
East
Lowry AFB
Denver, CO 80230

Dr. Arnold F. Kanarick
Honeywell, Inc.
2600 Ridgeway Pkwy.
Minneapolis, MN 55413

Dr. Milton S. Katz, Chief
Individual Training & Performance
Evaluation Technical Area
U.S. Army Research Institute
1300 Wilson Boulevard
Arlington, VA 22209

Dr. Steven W. Keele
Dept. of Psychology
University of Oregon
Eugene, OR 97403

Dr. David Klahr
Dept. of Psychology
Carnegie-Mellon University
Pittsburgh, PA 15213

Dr. Alma E. Lantz
University of Denver
Denver Research Institute
Industrial Economics Division
Denver, CO 80210

Dr. Sylvia R. Mayer (MCIT)
HQ Electronic System Division
LG Hanscom Field
Bedford, MA 01730

William J. McLaurin
Room 301
Internal Revenue Service
2221 Jefferson Davis Hwy.
Arlington, VA 22202

Dr. Donald E. Meyer
U.S. Air Force
ATC-XPTD
Randolph AFB, TX 78148

Military Assistant for Human Resources
Office of the Director of Defense
Research & Engineering
Room 3D129, The Pentagon
Washington, DC 20301

Dr. Andrew R. Molnar
Science Education Dev. & Res.
National Science Foundation
Washington, DC 20550

Dr. Ross L. Morgan (AFHRL-ASR)
Wright-Patterson AFB
Ohio 45433

Dr. Leo Munday
Houghton Mifflin Co.
P.O. Box 1970
Iowa City, IA 52240

Navy Personnel R&D Center
Code 306
San Diego, CA 92152
Attn: Dr. James McGrath

Navy Personnel R&D Center
San Diego, CA 92152
Attn: Dr. J.D. Fletcher

Navy Personnel R&D Center
Code 1
San Diego, CA 92152
Attn: Dr. James McGrath

Navy Personnel R&D Center
San Diego, CA 92152
Attn: Library

Dr. Leon Nawrocki
U.S. Army Research Institute
1300 Wilson Boulevard
Arlington, VA 22209

Dr. Donald A. Norman
Dept. of Psychology C-009
University of California, San Diego
La Jolla, CA 92093

Dr. Harold F. O'Neil, Jr.
Advanced Research Projects Agency
Cybernetics Technology, Room 623
1400 Wilson Blvd.
Arlington, VA 22209

Mr. Thomas C. O'Sullivan
TRAC
1220 Sunset Plaza Dr.
Los Angeles, CA 90069

ONR Branch Office
495 Summer Street
Boston, MA 2210
Attn: Dr. James Lester

ONR Branch Office
1030 East Green Street
Pasadena, CA 91101
Attn: Dr. Eugene Gloye

ONR Branch Office
536 S. Clark Street
Chicago, IL 60605
Attn: Dr. Charles E. Davis

Office of Naval Research
Code 200
Arlington, VA 22217

Mr. Luigi Petrullo
2431 N. Edgewood St.
Arlington, VA 22207

Dr. Kenneth A. Polycyn
PCR Information Sciences Co.
Communication Satellite Applications
7600 Old Springhouse Rd.
McLean, VA 22101

Principal Civilian Advisor
for Education and Training
Naval Training Command, Code 00A
Pensacola, FL 32508
Attn: Dr. William L. Maloy

R.Dir. M. Rauch
P II 4
Bundesministerium der Verteidigung
Postfach 161
53 Bonn 1, GERMANY

Research Branch
AFMPC-DPMYP
Randolph AFB, TX 78148

Dr. Joseph W. Rigney
University of So. California
Behavioral Technology Laboratories
3717 S. Grand
Los Angeles, CA 90007

Dr. Marty Rockway (AFHRL-TTI
Lowry AFB
Colorado 80230

Dr. Andrew M. Rose
American Institutes for Research 1055
Thomas Jefferson St. NW
Washington, DC 20007

Dr. Leonard L. Rosenbaum
Chairman
Dept. of Psychology
Montgomery College
Rockville, MD 20850

Dr. Worth Scanland
Chief of Naval Education & Training
NAS, Pensacola, FL 32508

Scientific Advisor to the Chief
of Naval Personnel (Pers Or)
Naval Bureau of Personnel
Room 4410, Arlington Annex
Washington, DC 20370

Dr. Robert J. Seidel
Instructional Technology Group, HumRRO
300 N. Washington St.
Alexandria, VA 22314

A.A. Sjoholm, Head Technical Support
Navy Personnel R&D Center
Code 201
San Diego, CA 92152

Dr. A. L. Slafkosky
Scientific Advisor (Code RD-1)
HQ, U.S. Marine Corps
Washington, DC 20380

Dr. Marshall S. Smith
Associate Director
NIE-OPEPA
National Institute of Education
Washington, DC 20208

Dr. Alfred F. Smode, Director
Training Analysis & Evaluation Group
Department of the Navy
Orlando, FL 32813

Dr. Richard Snow
Stanford University
School of Education
Stanford, CA 94305

LCDR J.W. Snyder, Jr.
F-14 Training Model Manager
VF-124
San Diego, CA 92025

Dr. William Strobie
McDonnell-Douglas Astronautics Co.
East
Lowry AFB
Denver, CO 80230

Dr. Persis Sturgis
Dept. of Psychology
California State University-Chico
Chico, CA 95926

Superintendent (Code 1424)
Naval Postgraduate School
Monterey, CA 93940

Technical Director
U.S. Army Research Institute for the
Behavioral & Social Sciences
1300 Wilson Boulevard
Arlington, VA 22209

Capt. Jack Thorpe, USAF
AFHRL-FTS
Williams AFB, AZ 85224

Mr. Walt W. Tornow
Control Data Corporation
Corporate Personnel Research
P.O. Box 0-HQN060
Minneapolis, MN 55440

Dr. Benton J. Underwood
Dept. of Psychology
Northwestern University
Evanston, IL 60201

Dr. Carl R. Vest
Battelle Memorial Institute
Washington Operations
2030 M Street NW
Washington, DC 20036

Dr. Joseph Ward
U.S. Army Research Institute
1300 Wilson Boulevard
Arlington, VA 22209

Dr. Claire E. Weinstein
Educational Psychology Dept.
University of Texas at Austin
Austin, TX 78712

Dr. David J. Weiss
Dept. of Psychology
N660 Elliott Hall
University of Minnesota
Minneapolis, MN 55455

Dr. Keith Wescourt
Dept. of Psychology
Stanford University
Stanford, CA 94305

Dr. Joseph L. Young
Director
Memory & Cognitive Processes
National Science Foundation
Washington, DC 20550

Robert Young
Advanced Research Projects Agency
1400 Wilson Blvd.
Arlington, VA 22209

122